This study investigated whether second language (L2) proficiency affects pragmatic comprehension, namely the ability to comprehend implied meaning in spoken dialogues, in terms of accuracy and speed of comprehension. Participants included 46 native English speakers at a U.S. university and 160 Japanese students of English in a college in Japan who were at different L2 proficiency levels. They took a 38-item computerized listening task measuring their ability to comprehend conversational implicatures of different types (i.e., 2 sets of items in different degrees of conventionality). The participants’ comprehension was analyzed in terms of accuracy (i.e., test scores on a multiple choice measure) and comprehension speed (i.e., average time in seconds taken to answer each item correctly). The results revealed a significant L2 proficiency influence on accuracy, but not on comprehension speed. There was no significant relationship between accuracy and comprehension speed. A post hoc analysis of error data revealed a short-term memory influence on comprehension accuracy for L2 learners.

THE EMERGENCE OF THE CONCEPT OF communicative competence (Bachman, 1990; Bachman & Palmer, 1996; Canale & Swain, 1980) marked a shift in the view of second language (L2) learning from mastery of grammatical forms alone to the acquisition of functional usage of forms in social contexts. According to this view, communicative competence includes pragmatic competence, namely the ability to comprehend and produce meaning in context.

What is important about meaning interpretation, from a pragmatic point of view, is understanding what the speaker intends to accomplish in making an utterance. Speakers make a variety of communicative choices, both verbal and non-verbal, which they enact for strategic reasons. The speaker’s choices may not always be transparent to the listener. Thus, it is an indispensable part of a listener’s pragmatic competence to recognize speaker intention conveyed in a variety of implicit and explicit forms and to make accurate inferences about a speaker’s communicative goals. Despite their importance, pragmatic inferential skills in L2 have been a relatively underrepresented area of investigation. Although abundant research has analyzed pragmatic production, particularly speech act performance, little research has examined L2 comprehension processes from a pragmatic perspective. Thus, the present research aims to fill the gap by analyzing L2 learners’ ability to comprehend implied meaning.

BACKGROUND

With the explicit recognition of the role of pragmatic competence in communicative ability, many L2 studies have analyzed learners’ pragmatic performance in communicative contexts. Over the past decade, we have seen how pragmatic development in L2 toward native-speaker norms positively correlates with overall L2 proficiency (Bardovi-Harlig & Dörnyei, 1998; Bardovi-Harlig & Hartford, 1993; Churchill, 2000; DuFon, 2001; Ellis, 1992; Hill, 1997; Roever, 2001; Rose, 2000; Taguchi, 1999; Trosborg, 1995). However, in the existing literature, pragmatic competence has been examined mainly based on production skills, specifically performance of speech acts (see e.g., Hill, 1997; House, 1996; Hudson, Detmer, & Brown, 1995; Maeshiba, Kasper, & Ross, 1996;
Pragmatic comprehension involves the ability to understand implied speaker intention by using linguistic knowledge, contextual clues, and the assumption of relevance (Grice, 1975; Levinson, 1983; Sperber & Wilson, 1995; Thomas, 1995). Relevance theory (Sperber & Wilson, 1995) argued that communication is never achieved by the mere decoding of linguistic stimuli; communication includes interpreting contextual cues and using them as evidence toward the correct inferencing of speaker intentions. When the speaker says something, the hearer assumes that the message is relevant to discourse context and thus seeks the most relevant and plausible interpretation of the intended meaning by using contextual information. Relevance theory defines context from both external and internal perspectives. The external context includes sociocultural parameters that index language use situations, such as physical settings and interlocutor relationships. The internal context, however, refers to the cognitive environment, including knowledge about the world, assumptions, beliefs, and experiences. Thus, inferential process is considered a global process because the hearer uses all conceptual information available in his or her cognitive environment.

Only a few L2 studies to date have examined pragmatic comprehension, particularly in the area of comprehension of implied meaning (Bouton, 1992, 1994; Carrell, 1981, 1984; Kasper, 1984; Koike, 1996; Taguchi, 2002; Takahashi & Roitblat, 1994; Ying, 1996, 2001). Bouton’s studies (1992, 1994) analyzed whether L2 learners could interpret implicatures by having L2 learners of English (N = 436) take a test that consisted of 33 short dialogues including implicatures. The results showed that, compared with that of native speakers, the learners’ performance in interpreting the implicatures was significantly poorer; however, learner comprehension improved over the 17-month period of their stay in the target country. In Bouton’s studies, language background and nationality of the L2 participants also affected comprehension, suggesting that interpretation of implicatures could vary cross-culturally.

In several other studies, L2 learners successfully identified the intended meaning of indirect speech acts (Kasper, 1984; Koike, 1996; Takahashi & Roitblat, 1994). Takahashi and Roitblat investigated L2 comprehension of conventionalized indirect requests. In their study, Japanese students with Test of English as a Foreign Language (TOEFL) scores over 500 (N = 36) and American students (N = 36) read 12 stories: 6 inducing implied requests and 6 inducing literal interpretations of the utterances. Results showed that, although the learners took longer to comprehend the indirect requests than did the native speakers, both groups arrived at the context-appropriate interpretations. Cook and Liddicoat (2002) investigated L2 comprehension of request speech acts at different levels of directness. In their study, 50 native speakers of English and 100 English as a second language (ESL) learners at two proficiency levels responded to a written questionnaire that contained 15 brief scenarios followed by a request-making expression. The request expressions had three directness levels: direct (e.g., “Please pass me the salt”), conventional indirect (e.g., “Can you pass me the salt?”), and nonconventional indirect (e.g., “Are you putting salt on my meat?”). Results showed that the native speakers were equally accurate at all directness levels. Although the high-proficiency learners had significant difficulty in understanding nonconventional indirect forms, the low-proficiency group showed difficulty in both conventional and nonconventional indirect forms. These results suggest that different types of indirect expressions make differential processing demands. Some expressions are less direct and less conventionalized than others and consequently increase the comprehension difficulty. The study also suggests that comprehension difficulty interacts with L2 proficiency: the higher their proficiency, the wider the range of directness levels that learners can handle.

Differential comprehension load across implied meaning types was also documented by the author (Taguchi, 2002). In this study, 8 Japanese learners of English listened to 15 dialogues in which the speakers’ final reply violated Grice’s (1975) relevance maxim and did not provide a direct response to the question. Following Holtgraves (1999), the 15 dialogues were divided into three categories: indirect refusals, indirect disclosure of information, and indirect expression of opinions. Analyses of the learners’ verbal protocols revealed that comprehension of indirect information disclosure and opinion expressions called upon a wider range of strategies, and elicited them more frequently than did comprehension of indirect refusals. The results suggest that comprehension of refusals requires
less effort. It is probable that the learners did not have to use many strategies to interpret indirect refusals because they were embedded in conventional, easily recognizable contexts.

The notion of different implied meaning types has been elaborated in several pragmatic theories. Relevance theory (Sperber & Wilson, 1995), for instance, emphasizes the relationship between context and processing effort. The theory claims that humans process information as productively as possible. When people interpret a message, many different assumptions from diverse sources come to mind. From the assumptions, they select the most relevant interpretation that has the greatest contextual effects for the smallest processing effort. Processing effort is reflected in the number and intensity of contextual signals to be interpreted (e.g., complexity of information, intonation, and nonverbal cues).

One factor that may reduce processing effort required for implicatures is the level of conventionality encoded in utterances. Conventionality was discussed earlier by Grice (1975) in the distinction between conventional implicatures and conversational implicatures, which are less conventional in nature. In conventional implicatures, the listener draws inferences from the utterance based on the conventional meanings encoded in lexical items. For example, when someone says, “She has two daughters,” the utterance is interpreted as “She has exactly two daughters,” but not as “She has at least two daughters.” Some indirect speech acts that have conventionalized linguistic forms (e.g., Would you mind + verb when making a request; Blum-Kulka, House, & Kasper, 1989) may share this nature of conventional implicatures. For these conventional implicatures, meaning is more conventional and context-independent, requiring less processing effort. In contrast, other less conventionalized implicatures, what Grice called conversational implicatures, are understood only when context factors are taken into account, as shown in the following example.

A: Has John left for work?
B: It’s after nine.

Speaker A can only understand B’s reply (i.e., whether John is still home or not) if he has contextual knowledge such as John’s routine and the time of the day, and if the questioner assumes that B has provided information relevant to the question. Thus, meaning decoding for this exchange is nonconventional and relies on specific contextual assumptions and consequently requires greater processing effort than would decoding a more conventional exchange.

In sum, different degrees of conventionality among implicatures are related to the differential processing effort required for meaning decoding. As Sperber and Wilson (1991, 1995) argued, implicatures vary in their degree of strength; some implicatures are strongly conveyed, while others are weakly understood, due to the number of contextual cues that must be processed for comprehension. The greater the number of cues to be processed, the more extensive the search for meaning becomes, resulting in greater processing effort. When implicatures convey conventional meaning, that is, when speaker intentions are linguistically coded or embedded within predictable, fixed patterns of discourse, the listener may not attend to such contextual cues as background knowledge, mutual understanding, or flow of discourse, consequently reducing the processing effort. The degree of indirectness or implicitness of an utterance is closely related to the amount of “work” that the listener has to do to comprehend meaning and is largely influenced by the amount of contextual information required for comprehension.

Despite these theoretical grounds, the use of pragmatic inferencing in L2 listening comprehension has not been widely studied and thus awaits research. In addition, the effect of conventionality in L2 pragmatic comprehension has not been studied systematically. Because very few previous studies have directly compared the comprehension of more and less conventionalized implicatures, integration of different implied meaning types into the design of a study could add to our understanding of pragmatic comprehension in a L2.

Another gap in previous research is that many studies have focused on the learner’s end product—how appropriately and accurately the learner can perform pragmatic functions—but very few studies have addressed the dimension of fluency or processing speed in learners’ pragmatic performance. Fluent processing deserves an independent analysis, apart from analysis of accuracy or appropriateness, because accuracy/appropriateness and speed are considered to form different dimensions of language performance (Brumfit, 2000; Koponen & Riggenbach, 2000; Schmidt, 1992). The attributed differences between speed and accuracy have also been documented in the field of cognitive psychology, in what is called the accuracy-speed trade off (Lachman, Lachman, & Butterfield, 1979; Oomen & Postma, 2001). In general, increasing
speed causes errors, and increasing accuracy costs time, as observed in both speech comprehension and production processes. Supporting this claim, in L2 performance studies, Skehan (1998) also argued that accuracy and speed entered into competition with one another; greater accuracy may be achieved at the expense of greater fluency, and vice versa.

In interlanguage pragmatics, fluency refers to automatic control in using pragmatic knowledge in real time (Kasper, 2001). Real-time comprehension involves a rapid transformation of information into thought, or the ability to attend to and understand quickly the intended interpretations in contexts. Speed, another dimension of comprehension, is influenced by the degree of inferencing involved in comprehension. Previous research documented that comprehension speed is usually faster when the old proposition is immediately accessible and slower when there is difficulty relating the current proposition to the preceding one (Just & Carpenter, 1980; Swinney, 1979; Till, Mross, & Kintsch, 1988). Comprehension of implied meaning is considered to take longer than that of conventional meaning because it is a type of perceived mismatch between old and new information. Decoding speed may vary among different types of implied meaning because of differential accessibility of information. As Rumelhart (1993) claimed, because of routinization of associative connections in long-term memory, more conventional utterances or contexts do not require extensive analytical procedures and thus are processed faster than less conventional ones.

Native language (L1) comprehension speed over different levels of pragmatic meaning was examined by Hamblin and Gibbs (2003). Native speakers of English read a series of seven sentences in which the final sentence conveyed either simple or complex pragmatic meaning. In the former case, what the speaker said and implied was similar (e.g., a response to the question about the noise level of an apartment was “This is a very noisy building”), whereas in the latter case, what the speaker said and implied was very different (e.g., a response to the same question about the noise level of an apartment was “I usually sleep with earplugs”). The results revealed that the participants took a significantly longer time to comprehend the dissimilar targets, suggesting that, while comprehending complex pragmatic meaning, the participants needed extra time to recognize multiple intended meanings and to identify the correct one. In the L1, comprehension of dissimilar meaning required more processing time because the participants used additional contextual effects and elaboration in meaning decoding.

Because few L2 studies have recorded processing time of nonliteral utterances, analyzing comprehension time could offer interesting insights into the interaction between overall proficiency and inferential ability. In L2, response time measures have been mostly used in psycholinguistic research to examine lower-order, microlevel processes only, such as lexical decisions, grammaticality judgments, and word-by-word reading (see e.g., Jiang, 2002; Juffs, 1998; Neumann, McLoskey, & Felio; 1999; Segalowitz, Segalowitz, & Wood, 1998; Snellings, van Gelderen, & Gloppe, 2002; White & Juffs, 1998). The results generally suggest that learners’ processing fluency is associated with automatization of lower-order processes and that such automaticity is related to language proficiency (Carver, 1990; Hirai, 1999; Segalowitz, 2001; Towell, 2002; White & Genesee, 1996).

Only a few studies have examined the relationship between speed of lower-level processing and higher-order processing (e.g., reading comprehension of authors’ intentions or relationship among paragraphs; Fukkink, Hulstijn, & Simis, 2005). Very little research has investigated processing speed of implied intentions in listening. Thus, it would be interesting to find out whether previous findings about lower-order processing alone can be extended to higher-level processing, such as processing of implied meaning, in listening skill. Such information will show whether L2 pragmatic competence can also be characterized by learners’ increased speed in comprehending implied meaning.

In sum, L2 pragmatic performance in listening comprehension has been a relatively underrepresented area of investigation. In addition, issues related to the comprehension of different types of implied meaning need to be examined for learners of different proficiency levels. Such an investigation could reveal whether previous findings about the positive relationship between L2 proficiency and pragmatic production can be extended to the area of pragmatic comprehension. Finally, comprehension speed could well serve as an additional useful tool to characterize developmental aspects of pragmatic knowledge. Whereas information on accuracy of comprehension would support inferences about learners’ pragmatic knowledge, information on speed of comprehension would support inferences about the processing of pragmatic knowledge. Data on accuracy and speed would provide complementary information about learners’ pragmatic competence and might consequently guide us toward
better understanding of learners’ communicative competence. With increasing attention in L2 research to the relationship between basic cognitive abilities and language performance, researchers believe that cognitive systems such as attention allocation, memory, and speedy lexical access skills affect performance in language-related tasks (Robinson, 2003; Segalowitz & Freed, 2004). Thus, L2 development is characterized in part by the improved operations of cognitive processes while accessing L2 knowledge, as manifested in speed of response. Based on this rationale, this study examined two attributes of pragmatic knowledge during comprehension: accuracy and response speed.

RESEARCH QUESTIONS

This study aimed to examine L2 comprehension of different types of implied meaning and to compare learners’ performance with native speaker performance. The study examined how L2 learners of English comprehend different types of implied meaning in terms of accuracy and speed of comprehension. The study also investigated whether L2 proficiency predicted accuracy and speed of comprehension of different types of implied meaning. Finally, the study sought a relationship between accuracy and speed in L2 pragmatic comprehension. The following three research questions guided the study:

1. Do L2 learners differ in accuracy and comprehension speed for different types of implied meaning?
2. Does L2 proficiency affect accuracy or comprehension speed in response to different types of implied meaning?
3. Is there a relationship between L2 learners’ accuracy and comprehension speed?

OPERATIONALIZING THE CONSTRUCT OF PRAGMATIC COMPREHENSION

Through the mode of listening, this study measured L2 learners’ ability to understand a speaker’s implied meaning in conversations accurately and rapidly. While completing the task, the L2 learners played the role of “overhearer” of the conversations; they listened to the dialogues and identified implied meaning encoded in the dialogues they heard.

In the field of cognitive psychology, comprehension is defined as a process of receiving, attending to, and assigning meaning to aural/visual stimuli (Anderson, 1990; Wolvin & Coakley, 1985). The comprehension process is considered interactive, with bottom-up and top-down processing operating in parallel manners (McClelland & Rumelhart, 1981). Letters or words perceived through bottom-up processing are assigned meaning through feedback from a higher level, such as contextual cues or prior knowledge. Smith (1975) also claimed that, in order to make sense of what is said, acoustic information must be supplemented with nonacoustic information, such as knowledge of the world, language, and context. Pragmatic knowledge can be categorized as nonacoustic information and part of top-down processing, guiding the interpretation of the lower-level input. Successful interpretation of speaker intention, encoded in acoustic information, depends on the use of pragmatic cues, such as contextual information, schemata, or knowledge of communication.

Through the process of listening, the participants in this study derived the inferences necessary to select responses on an objective test. They engaged in a series of inferential processes: interpreting the language and constructing a representation of the conversational meaning; constructing an interpretation of the test item stem; understanding how to relate the memory of the conversation with the test response choices given; and, finally, making the best selection, based on understanding of the task. These inferential processes were examined to see how they contributed to the comprehension of implied meaning.

Implied meaning is defined as meaning “that goes beyond what is given by the language form itself or what is literally said” (Verschueren, 1999, p. 25). Thus, comprehension of implied meaning refers to the ability to recognize a mismatch between the literal utterance and the intention of the utterance and to comprehend the intention of the utterance. Two types of implied meaning were tested in the present study: more conventional implicatures (MCI) and less conventional implicatures (LCI). In this study, the item difference lay in the amount of inferencing required for comprehension affected by the degree of conventionality. MCI items in this study included specific types of indirect requests and refusals. They were more conventionalized forms of implied meaning because specific linguistic forms or language use patterns were available, as suggested by the previous literature (Gibbs, 1983; Morgan, 1978). For instance, the linguistic form Do you mind followed by a verb usually appears when one is making an indirect request (Blum-Kulka et al., 1989). When making indirect refusals, it is a conventional language use pattern for one to provide a reason
for refusing (Beebe, Takahashi, & Uliss-Weltz, 1990; Nelson, Carson, Batal, & Bakary, 2002). These two examples of requests and refusals parallel the two types of conventions proposed by Morgan (1978): conventions of language and conventions of usage. In the request example, conventionality is encoded in a specific linguistic form whereas in the refusals, conventionality appears when certain expressions are customarily used.

In contrast to the MCI items, the LCI items in this study were defined as a means to provide information in an indirect manner. In the LCI items, meaning was not simply attached to specific linguistic expressions or uses. For instance, in the following example, there are many different ways to convey a positive opinion of one’s neighbors.

A: Do you like your neighbors upstairs?
B: We’re always visiting each other.

Thus, compared with the MCI expressions, those used for LCI were more idiosyncratic and less stable across language users and contexts. It was assumed that listeners would need to use greater inferential processing to understand their meanings than the meanings of the MCI expressions because the listeners would process a greater number of contextual cues (e.g., background knowledge, flow of discourse).

METHODOLOGY

Participants

The study included 210 participants, consisting of two distinct groups: 46 native speakers of English in a university in the United States, and 164 Japanese learners of English in a branch U.S. college in Japan. The 46 native English speakers (20 males and 26 females) were volunteer participants enrolled in undergraduate courses at a medium-sized university in the southwest United States. Their ages ranged from 19 to 31 with a mean age of 22.5 years. Of the Japanese participants, 98 came from the Intensive English Program (IEP) in the school, whereas the remaining 66 came from the school’s General Education (GE) division. Four Japanese students who had extensive living experience in English-speaking countries (i.e., more than 1 year) were excluded from the sample. As a result, a total of 160 Japanese participants were used for the final analyses, 81 males and 79 females. Their average age was 20.5 years, ranging from 17 to 36. The Japanese students in this study were exposed to spoken English mainly through their classes because English was the instructional medium in the branch U.S. university. Because approximately 90% of the student population was Japanese, and all of them lived in Japan, the students did not have extensive L2 English input outside the class. Table 1 summarizes information about the Japanese students’ general L2 English proficiency measured by the institutional TOEFL.

Instrumentation

The participants’ ability to comprehend implied meaning was assessed by a pragmatic listening task, consisting of 40 multiple choice questions: 2 practice items, 16 MCIs, 16 LCIs, and 6 filler items. In addition to the target implicature items, filler items that tested literal comprehension were included in the task in order to divert the participants’ attention from the purpose of the experiment, namely testing comprehension of nonliteral, implied meaning. The task was computerized using PsyScope 1.2.5. (Cohen, MacWhinney, Flatt, & Provost, 1993).

In most existing L2 research, comprehension ability of implied meaning has been measured by a reading instrument (i.e., participants read conversations and identify implied meaning). However, this study tested the same ability with a listening task in order to increase the authenticity of the task and contribute to construct validity. According to Bachman and Palmer (1996), one type of evidence needed for validity is the degree of correspondence between test tasks and the target language use (TLU) domain, where learners’ ability can be generalized beyond the testing

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Institutional TOEFL Test Descriptive Statistics</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Composite Score</td>
</tr>
<tr>
<td>Listening Section Score</td>
</tr>
</tbody>
</table>

Note. The descriptive statistics are based on the 160 Japanese L2 learners of English who took the institutional TOEFL. Composite score includes scores from the subsections of listening, grammar, and reading.
situation to the real life situation. In the TLU domain, conversations are usually heard, not read, and also require processing; listeners cannot control the rate at which the information is taken in. Thus, a listening task is more likely than a reading task to elicit an adequate representative sample of the TLU domain. In order to increase the degree of correspondence between test tasks and TLU domain, dialogues for this study were written to sound as natural as possible. Consequently, linguistic units that characterize the interactive nature of spoken English, such as discourse markers (e.g., well, you know), interjections (e.g., oh), or hesitation markers (e.g., um; see Biber, Johansson, Leech, Conrad, & Finegan, 1999), were included in the dialogues.

Tables 2 and 3 display simplified sample items. MCI items included indirect requests and refusals. The indirect requests included fixed linguistic forms listed under “conventional indirect” in the Cross Cultural Speech Act Realization Project (CCSARP; see Blum-Kulka et al., 1989), including forms such as I was wondering if you could + VP (verb phrase), Do you mind if I + VP, and Do you think you can + VP. In addition, a few items included expressions of “need” (e.g., We need some food too, don’t you think?) and indications of a problem (e.g., Your music is loud, and I’m trying to get some sleep), which were identified as conventional request expressions in the previous literature (Davis, 1998). Indirect refusals, however, were written so as to avoid explicit linguistic markers of refusals such as No, I can’t, or I don’t want to, which were identified as direct refusal expressions by Beebe, et al. (1990). Utterances used for indirect refusals took the form of providing a reason.

Unlike the MCI items, the LCI items for this study were written as expressions used to convey opinions or to provide some information indirectly without specific conventionalized features (i.e., fixed linguistic expressions or conventions of language use). In order to counteract any response biases, half of the LCI items and distractors were designed to exemplify a negative intended meaning (i.e., the utterances conveyed negative opinions or information), whereas the other half of the LCI items were written to imply nonnegative meaning (i.e., the utterances conveyed either positive or neutral opinions or information). These design features were incorporated to add to the breadth of conclusions that the study and its results may justify.

For item writing, the researcher consulted six references in the field of comprehension of implicatures: Bouton (1992, 1994), Holtgraves (1999),

---

### TABLE 2
Simplified Sample More Conventional Implicature Items

<table>
<thead>
<tr>
<th>Indirect Requests</th>
<th>Indirect Refusals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jane: Do we have time to go over my paper?</td>
<td>Mary: Let’s go to the movies tonight.</td>
</tr>
<tr>
<td>Dr. White: Oh, ah, do you mind if we talk about it tomorrow?</td>
<td>John: I have to finish my paper by eight in the morning.</td>
</tr>
<tr>
<td><strong>Q:</strong> Which of the following is correct?</td>
<td><strong>Q:</strong> Which of the following is correct?</td>
</tr>
<tr>
<td>(a) Dr. White doesn’t like Jane’s paper.</td>
<td>(a) John wants to go to the movies.</td>
</tr>
<tr>
<td>(b) Dr. White doesn’t want to go over the paper.</td>
<td>(b) John is going to take a break.</td>
</tr>
<tr>
<td>(c) Dr. White is asking Jane to come back tomorrow.</td>
<td>(c) John has to finish his paper by 8:00 tonight.</td>
</tr>
<tr>
<td>(d) Dr. White is asking Jane to help his work.</td>
<td>(d) John can’t go to the movies tonight.</td>
</tr>
</tbody>
</table>

---

### TABLE 3
Simplified Sample Less Conventional Implicature Items

<table>
<thead>
<tr>
<th>Negative Implied Meaning</th>
<th>Nonnegative Implied Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>John: How was the wedding? I bet it was exciting.</td>
<td>Dave: Do you like the people upstairs?</td>
</tr>
<tr>
<td>Mary: Well . . . the cake was OK.</td>
<td>Susan: We’re always visiting each other.</td>
</tr>
<tr>
<td><strong>Q:</strong> What does Mary think about the wedding?</td>
<td><strong>Q:</strong> What does Susan think about the people living upstairs?</td>
</tr>
<tr>
<td>(a) Mary doesn’t remember the wedding.</td>
<td>(a) She likes them very much.</td>
</tr>
<tr>
<td>(b) Mary thought that the cake wasn’t so good.</td>
<td>(b) She doesn’t know them so well.</td>
</tr>
<tr>
<td>(c) Mary liked the wedding very much.</td>
<td>(c) She doesn’t like them so much.</td>
</tr>
<tr>
<td>(d) Mary didn’t enjoy the wedding very much.</td>
<td>(d) She has never visited them.</td>
</tr>
</tbody>
</table>
Kotsonis (1981), Roever (2001), and Takahashi and Roitblat (1994). Several items were adapted from these sources with the authors’ permissions. The conversations, questions, and four answer options were written by the researcher with help from a native English speaker who had experience with English as a foreign language (EFL) students. When writing the dialogues, great care was taken to keep the lengths of all of the conversations approximately the same in order to control impact on short-term memory. Each conversation had between 42 and 47 words, with a mean of 44.4 and standard deviation of 1.9 words. The number of words in question and option sentences was also kept equal across item categories in order to reduce the effect of extraneous variables (e.g., reading time) and make response times comparable. All of the questions that appeared in the task took either the form of a Wh-question (e.g., What did she think about the movie?) or of a more general question type (e.g., Which of the following statements is correct?). The number of these two question formats was balanced across item categories.

Furthermore, following Roever (2001), in order to reduce the effect of construct-irrelevant variance from the learners’ difference in vocabulary knowledge, all vocabulary in the items was drawn from Longman’s 2,000-word defining vocabulary list (Longman, 1995). The 2,000 words on the list are identified as common, frequent, and basic English words and are thus used to write all the word definitions in the Longman dictionary. Thus, the 2,000-word-level vocabulary items were considered relatively attainable by the L2 participants.

In order to confirm the item classifications as either MCI or LCI, each conversation was coded by a second rater, a native speaker of English who had knowledge of pragmatics. The agreement rate between the two raters was .97. The one item that the two raters did not agree on was revised. The test was piloted with two adult native speakers of English, both of whom answered 100% of the items correctly. Several items were revised based on their feedback.

Each listening item followed a multiple choice question format with one correct answer and three distractor options. The focus of investigation was to examine whether the participants were able to identify the correct implied intention, without being distracted by other answer choices that were not correct, but were related to overall conversation. The three principles below were applied when writing the distractors in 20 of the 32 implied meaning items:

Principle 1: The option contains a meaning that is the opposite of the implied meaning.
Principle 2: The option contains words taken from the last part of the dialogue.
Principle 3: The option is related to the overall conversation.

The example below illustrates how these principles were applied:

Ben: Morning honey. I can’t believe I fell asleep in the middle of the movie last night. Did you watch it till the end?
Barbara: Yeah, I did.
Ben: How was it? Did you like it?
Barbara: I was glad when it was over.

Question: What did Barbara think about the movie?

(a) She thought the movie was good. (Principle 1)
(b) She didn’t enjoy the movie. (Correct answer)
(c) She thinks Ben should have watched the movie. (Principle 3)
(d) She liked the end of the movie. (Principle 2)

The pragmatic listening task was piloted with 58 participants (ESL learners and native speakers). Further modifications were made based on the results of the pilot study. In the main study, the internal consistency reliability of the listening task was estimated using the Cronbach’s alpha, yielding .92 for the full test, .83 for the MCIs (16 items), and .85 for the LCIs (16 items).

Data Collection Procedures

Data collection took place in the summer of 2002, following approval of the study by the Institutional Review Board at the researcher’s home university. The listening task was given using Macintosh computers during class periods. After being individually seated in front of computers, the participants put on headphones and read instructions on the screen in Japanese, their native language. Before starting the test items, they practiced on two items. Immediately following each conversation, a multiple choice question appeared on the screen. The participants chose the most appropriate response by pressing the corresponding number on the keyboard. After they chose an answer, the computer automatically took them to the next item. Response time was measured between the moment when the question appeared on the screen and the moment when the participants pressed the number key for their
answer choice. The computer recorded all responses and their latencies. The participants were not informed of the objective of the experiment and were told to proceed at their own pace. Approximately 3 weeks after the completion of the listening task, all Japanese participants took the institutional TOEFL on campus.

Data Analysis Procedures

This study examined whether accuracy and comprehension speed would differ according to L2 proficiency and the types of implied meaning tested. L2 proficiency, the independent variable in the study, was determined by the institutional TOEFL scores. Table 1 displays descriptive statistics of the composite TOEFL results and the listening section scores. The L2 participants had a wide range of general L2 proficiency, with institutional TOEFL scores ranging from 330, the beginner level, to 590, the high-intermediate level. The institutional TOEFL scores showed a moderately high correlation with the participants’ level of study in the institution (r = .76, p < .000), which was determined on the basis of the results of in-house proficiency exams each semester.

The institutional TOEFL scores, the independent variable, were treated as a continuous variable, ranging from 330 to 590 (ETS, 2001). L2 learners’ comprehension ability was operationalized in terms of accuracy of comprehension of two item types (i.e., MCI and LCI) and speed of comprehension of these two item categories. Accuracy was measured by the pragmatic listening task that had an interval scale between 0 and 16 for each item category. Comprehension speed, also interval data, was calculated by averaging the number of seconds taken to answer items correctly. These four observations (i.e., scores and times for MCIs and LCIs) formed four dependent variables. The influence of L2 proficiency on the four dependent variables was examined by using four separate regression analyses. Differences in accuracy and speed between MCIs and LCIs were examined using a matched pair t test. The relationship between accuracy and speed was addressed by partial correlation, after removing potentially irrelevant variance coming from the third variable, L2 proficiency (i.e., institutional TOEFL scores). The six filler items in the listening instrument were excluded from all analyses.

Prior to the statistical analyses, distributions of variables (i.e., the TOEFL scores, scores and comprehension times of MCIs and LCIs) were examined to check underlying assumptions. Data screening results documented that the distribution of time data did not meet the underlying assumptions of normality, linearity, and homoscedasticity. Because violation of these assumptions could weaken the statistical analysis, some statistical procedures to remedy these problems were necessary prior to the fundamental analysis.

When nonnormality is found, data transformation is recommended (Howell, 1997; Tabachnick & Fidell, 2001) in order to improve normality. Because the distribution was positively skewed, a logarithmic transformation was performed following Tabachnick and Fidell (2001). After the transformation, the distributions were found to be normal and linear, and homoscedasticity was confirmed. Significance tests for skewness and kurtosis (i.e., skewness and kurtosis values divided by their standard error) confirmed normality (p < .001). The Kolmogorov-Smirnov test also confirmed normality at the alpha level of .05. Thus, it was decided that the statistical analyses should be performed on the logarithmically transformed data. Based on previous conventions, the significance level for hypothesis testing was set at .05; however, because the study used six statistical comparisons, the alpha level was adjusted to .01 using the Bonferroni correction in order to avoid a Type I error (i.e., the error of rejecting the null hypothesis when it is true) (Hatch & Lazaraton, 1991; Jaeger, 1993; SPSS, 1998).

RESULTS

Research Question 1: Influence of Implied Meaning Types on Comprehension

The first research question in this study examined whether L2 learners differ in accuracy and comprehension speed shown for different types of implied meaning (i.e., MCIs and LCIs). Tables 4 and 5 display descriptive statistics of the scores and response times when comprehending MCIs and LCIs.

The statistics show that native speaker comprehension was accurate and generally fast, and little variation was observed in score and time data (i.e., SDs of 0.69 for scores and 2.18 for time), as compared to the statistics for L2 learners (i.e., SDs of 6.52 for scores and 6.01 for time). The results indicate that native speaker performance was relatively uniform in accuracy and processing speed. In addition, in native speaker comprehension, the descriptive statistics showed almost no difference in accuracy and comprehension speed between MCIs and LCIs. However, in L2 comprehension,
TABLE 4
Descriptive Statistics for Comprehension Scores

<table>
<thead>
<tr>
<th>Group</th>
<th>K</th>
<th>M</th>
<th>Mdn</th>
<th>SD</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSs Total</td>
<td>32</td>
<td>37.42</td>
<td>38.00</td>
<td>0.69</td>
<td>35.00</td>
<td>38.00</td>
</tr>
<tr>
<td>(n = 46)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCI</td>
<td>16</td>
<td>15.63</td>
<td>16.00</td>
<td>0.61</td>
<td>14.00</td>
<td>16.00</td>
</tr>
<tr>
<td>Refusals</td>
<td>8</td>
<td>7.91</td>
<td>8.00</td>
<td>0.28</td>
<td>7.00</td>
<td>8.00</td>
</tr>
<tr>
<td>Requests</td>
<td>8</td>
<td>7.76</td>
<td>8.00</td>
<td>0.48</td>
<td>6.00</td>
<td>8.00</td>
</tr>
<tr>
<td>LCI</td>
<td>16</td>
<td>15.67</td>
<td>16.00</td>
<td>0.59</td>
<td>14.00</td>
<td>16.00</td>
</tr>
<tr>
<td>Negative Meaning</td>
<td>8</td>
<td>7.93</td>
<td>8.00</td>
<td>0.25</td>
<td>7.00</td>
<td>8.00</td>
</tr>
<tr>
<td>Nonnegative Meaning</td>
<td>8</td>
<td>7.74</td>
<td>8.00</td>
<td>0.57</td>
<td>6.00</td>
<td>8.00</td>
</tr>
<tr>
<td>L2 Learners Total</td>
<td>32</td>
<td>25.54</td>
<td>26.00</td>
<td>6.52</td>
<td>5.00</td>
<td>37.00</td>
</tr>
<tr>
<td>(n = 160)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCI</td>
<td>16</td>
<td>12.11</td>
<td>13.00</td>
<td>3.06</td>
<td>2.00</td>
<td>16.00</td>
</tr>
<tr>
<td>Refusals</td>
<td>8</td>
<td>5.64</td>
<td>6.00</td>
<td>1.67</td>
<td>1.00</td>
<td>8.00</td>
</tr>
<tr>
<td>Requests</td>
<td>8</td>
<td>6.27</td>
<td>7.00</td>
<td>1.85</td>
<td>0.00</td>
<td>8.00</td>
</tr>
<tr>
<td>LCI</td>
<td>16</td>
<td>8.59</td>
<td>9.00</td>
<td>3.07</td>
<td>2.00</td>
<td>16.00</td>
</tr>
<tr>
<td>Negative Meaning</td>
<td>8</td>
<td>4.02</td>
<td>4.00</td>
<td>2.03</td>
<td>0.00</td>
<td>8.00</td>
</tr>
<tr>
<td>Nonnegative Meaning</td>
<td>8</td>
<td>4.58</td>
<td>5.00</td>
<td>1.55</td>
<td>1.00</td>
<td>8.00</td>
</tr>
</tbody>
</table>

Note. MCI = more conventional implicatures; LCI = less conventional implicatures. K refers to the number of items in each category. One point was given per correct answer. The highest possible score in each item category was 16.

TABLE 5
Descriptive Statistics for Average Response Time in Seconds

<table>
<thead>
<tr>
<th>Group</th>
<th>K</th>
<th>M</th>
<th>Mdn</th>
<th>SD</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSs Total</td>
<td>32</td>
<td>9.54</td>
<td>8.70</td>
<td>2.18</td>
<td>5.67</td>
<td>14.32</td>
</tr>
<tr>
<td>(n = 46)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCI</td>
<td>16</td>
<td>9.70</td>
<td>9.41</td>
<td>2.89</td>
<td>5.88</td>
<td>17.81</td>
</tr>
<tr>
<td>Refusals</td>
<td>8</td>
<td>10.44</td>
<td>9.65</td>
<td>3.05</td>
<td>6.07</td>
<td>18.93</td>
</tr>
<tr>
<td>Requests</td>
<td>8</td>
<td>9.01</td>
<td>8.34</td>
<td>2.96</td>
<td>4.82</td>
<td>18.50</td>
</tr>
<tr>
<td>LCI</td>
<td>16</td>
<td>9.57</td>
<td>8.78</td>
<td>3.19</td>
<td>5.07</td>
<td>20.06</td>
</tr>
<tr>
<td>Negative Meaning</td>
<td>8</td>
<td>7.91</td>
<td>7.11</td>
<td>2.25</td>
<td>3.84</td>
<td>12.84</td>
</tr>
<tr>
<td>Nonnegative Meaning</td>
<td>8</td>
<td>11.19</td>
<td>9.53</td>
<td>4.87</td>
<td>6.14</td>
<td>31.66</td>
</tr>
<tr>
<td>L2 Learners Total</td>
<td>32</td>
<td>19.97</td>
<td>18.94</td>
<td>6.01</td>
<td>9.64</td>
<td>44.54</td>
</tr>
<tr>
<td>(n = 160)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCI</td>
<td>16</td>
<td>19.31</td>
<td>18.89</td>
<td>6.33</td>
<td>9.27</td>
<td>48.58</td>
</tr>
<tr>
<td>Requests</td>
<td>8</td>
<td>22.25</td>
<td>19.58</td>
<td>10.28</td>
<td>9.52</td>
<td>88.84</td>
</tr>
<tr>
<td>Refusals</td>
<td>8</td>
<td>17.44</td>
<td>16.54</td>
<td>6.56</td>
<td>7.44</td>
<td>62.03</td>
</tr>
<tr>
<td>LCI</td>
<td>16</td>
<td>22.29</td>
<td>20.93</td>
<td>7.76</td>
<td>9.60</td>
<td>56.06</td>
</tr>
<tr>
<td>Negative Meaning</td>
<td>8</td>
<td>22.23</td>
<td>20.04</td>
<td>10.71</td>
<td>9.60</td>
<td>80.33</td>
</tr>
<tr>
<td>Nonnegative Meaning</td>
<td>8</td>
<td>23.49</td>
<td>21.00</td>
<td>11.85</td>
<td>2.51</td>
<td>92.14</td>
</tr>
</tbody>
</table>

Note. MCI = more conventional implicatures; LCI = less conventional implicatures. Comprehension time in seconds refers to the average number of seconds taken to comprehend each item correctly.

The mean response time was longer and mean scores were lower for the LCIs.

Comprehension scores and response time of all L2 participants (N = 160) were compared between the two item categories (i.e., MCI and LCI) by using a matched pair t test (α = .01) in order to see whether the observed difference in descriptive statistics was statistically significant. As shown in Table 6, the t test revealed a significant difference in comprehension accuracy, t(159) = 17.05, p < .01. The obtained effect size of 1.15 based on Cohen's d shows a very strong effect (Cohen, 1988). There was also a significant difference in comprehension speed, t(159) = 6.45, p < .01, with a moderate effect size of 0.43. The results supported the hypotheses; MCI items were significantly easier and took less time for L2 learners to comprehend than LCI items.

Research Question 2: Influence of L2 Proficiency on Comprehension

The second research question asked whether L2 proficiency affects accuracy or comprehension...
TABLE 6
Matched Pair t test Results for Accuracy and Speed: Comparison between More Conventional Implicatures and Less Conventional Implicatures
(N = 160)

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>t</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scores (Accuracy)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>More Conventional Implicatures</td>
<td>12.11</td>
<td>3.06</td>
<td>17.05*</td>
<td>1.15</td>
</tr>
<tr>
<td>Less Conventional Implicatures</td>
<td>8.59</td>
<td>3.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Comprehension Time by Seconds (Speed)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>More Conventional Implicatures</td>
<td>19.31</td>
<td>6.33</td>
<td>6.45*</td>
<td>0.43</td>
</tr>
<tr>
<td>Less Conventional Implicatures</td>
<td>22.23</td>
<td>7.76</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. K refers to the number of items. Effect size is based on Cohen’s d. *p < .01.

speed in response to different types of implied meaning. Regression analysis was used to assess the relationship between the independent variable, L2 proficiency, and the four dependent variables, the scores and response time by seconds in two implied meaning categories (i.e., MCI and LCI).

As displayed in Table 7, L2 proficiency, operationalized as institutional TOEFL scores (range: 330–590), was found to be a real predictor for the three of the dependent variables tested: scores of MCI, scores of LCI, and comprehension time of MCI, at the significance level of .01. However, there was no significant impact of L2 proficiency on the comprehension time of LCI at the significance level of .01, indicating that proficiency was not a predictor for comprehension speed of this type of implied meaning. The magnitude of coefficient was also very small (i.e., R-square = .03), suggesting that there was little change in comprehension time for LCIs across the institutional TOEFL scores. Another time variable, comprehension time for MCIs, was found to be predicted by L2 proficiency at the significance level of .01; however, the relatively small correlation coefficient (i.e., R-square = .11) indicates that this variable had a very weak relationship with L2 proficiency in this L2 participant group. The other two accuracy variables had higher correlation coefficients. The R-square of .41 for the accuracy of LCI shows that, of the original 100% of the variance, 41% of the variance was accurately accounted for by L2 proficiency. Similarly, for the accurate comprehension of LCI, 32% of the variance was explained by L2 proficiency. These results suggest that there was indeed a moderate to moderately strong effect of proficiency on accuracy, but not on response time.

TABLE 7
Regression Analysis of Proficiency Effect on Comprehension of Implied Meaning
(N = 160)

<table>
<thead>
<tr>
<th></th>
<th>Standardized Coefficients (Beta)</th>
<th>R Square</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>More Conventional Implicatures</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scores</td>
<td>0.57</td>
<td>0.32</td>
<td>8.61*</td>
<td>.000</td>
</tr>
<tr>
<td>Time</td>
<td>0.33</td>
<td>0.11</td>
<td>4.32*</td>
<td>.000</td>
</tr>
<tr>
<td>Less Conventional Implicatures</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scores</td>
<td>0.64</td>
<td>0.41</td>
<td>10.42*</td>
<td>.000</td>
</tr>
<tr>
<td>Time</td>
<td>0.17</td>
<td>0.03</td>
<td>2.17</td>
<td>.032</td>
</tr>
</tbody>
</table>

Note. Proficiency was based on learners’ composite scores of the institutional TOEFL, ranging between 330 and 590. *p < .01.

Research Question 3: Relationship between Accuracy and Comprehension Speed

The last research question asked whether there is a relationship between learners’ accuracy and speed of comprehension. Partial correlation was applied in order to reveal the relationship between accuracy and speed after removing the effect of the third variable, L2 proficiency. Partial correlation was considered appropriate to use because relatively high Pearson correlations were found between L2 proficiency, determined by institutional TOEFL scores, and accuracy of comprehension (i.e., .57 for MCI and .64 for LCI). Controlling for the effect of L2 proficiency, a Pearson correlation was calculated for all L2 learners between the total scores and average comprehension time by seconds for all the implied meaning items (MCI and LCI combined). The results of the partial correlation revealed no significant
relationship between comprehension scores and time, \( r = -0.14, p > .01 \).

**Post Hoc Analysis of Error Data in L2 Pragmatic Comprehension**

In order to gain some insights into L2 learners’ cognitive processes when responding to the listening items, two types of post hoc analysis of error data were conducted: examination of response time for correct versus incorrect items and distractor analysis for incorrect items. Table 8 summarizes the results of the first post hoc analysis. The average response time for wrongly answered items was longer than the time taken for correctly answered items for both item categories; furthermore, the standard deviation was also greater for erroneous items. The results suggest that there was greater variance in processing speed when the L2 learners drew wrong inferences, indicating possible extraneous factors at work when the learners were comprehending error items (e.g., lapses of attention, misreading sentences). It is also possible that the learners got stuck on particular items, resulting in slow response time. The results revealed by the partial correlation further confirmed the findings of no significant relationship between comprehension accuracy and speed. The speed-accuracy trade-off found in previous L1 research (Lachman et al., 1979; Oomen & Postma, 2001) was not evident in the present study, because errors did not result from fast comprehension speed.

The second post hoc analysis of error data reported analyses of response options for each multiple choice question in the pragmatic listening task. The three distractors for each item were established as (a) the option contained the opposite meaning to the implied meaning, (b) the option contained the exact words (or synonyms) taken from the last-heard utterance of the dialogue, and (c) the option contained words or content related to the overall conversation. Distractor analyses were performed on two groups of L2 learners of different proficiency. The 160 learners were divided into three groups according to their institutional TOEFL scores. The top third \((n = 51)\) and bottom third \((n = 54)\) were selected for comparison and were labeled as the higher- and lower-proficiency groups, respectively.

Table 9 displays the results of the distractor analysis. As shown in the table, the L2 learners were most attracted to the second type of distractor (i.e., words taken from the last-heard utterance). Regardless of the proficiency level, more than half of the incorrect answers were this distractor type. This recency effect became more discernible as item difficulty increased. In this study, LCIs were

| TABLE 8 | L2 Response Time (in Seconds) for Correct and Incorrect Implied Meaning Items \((N = 160)\) |
| --- | --- | --- | --- | --- | --- |
| | \( M \) | \( \text{Min} \) | \( SD \) | \( \text{Min} \) | \( \text{Max} \) |
| More Conventional Implicatures \((K = 16)\) | | | | | |
| Response Time, Correct Items | 19.31 | 18.89 | 6.33 | 9.27 | 48.58 |
| Response Time, Incorrect Items | 25.09 | 23.12 | 11.82 | 21.89 | 72.52 |
| Less Conventional Implicatures \((K = 16)\) | | | | | |
| Response Time, Correct Items | 22.29 | 20.93 | 7.76 | 9.60 | 56.06 |
| Response Time, Incorrect Items | 27.82 | 25.87 | 10.67 | 10.21 | 65.80 |

Note. \(K = \) number of items.

| TABLE 9 | Distractor Analysis for L2 Pragmatic Listening Task Data |
| --- | --- | --- | --- | --- |
| | Correct Responses \(\% (n)\) | Opposite to the Implicated Meaning \(\% (n)\) | Words from Last-Heard Utterance \(\% (n)\) | Related to Other Parts of Dialogue \(\% (n)\) |
| Higher-Proficiency L2 Group | 76.9 (784) | 5.5 (56) | 12.5 (128) | 5.1 (52) |
| Lower-Proficiency L2 Group | 50.0 (540) | 11.2 (121) | 27.5 (297) | 11.3 (122) |

Note. The results were based on the 20 items. There were 51 students in the higher L2 group and 54 students in the lower L2 group. Numbers in the table show percentages for each option chosen by the group. Numbers in the parentheses \(n\) show the raw counts.
found harder to comprehend than MCIs. The mean score for LCIs was 8.59, whereas for MCIs it was 12.10 out of 16 points possible \((N = 160)\). The distractor analysis showed that, when comprehending LCIs, more learners chose the distractor that contained the words taken from the last utterance they heard than the other distractors. For example, item 30 (Speaker A asks “Did you like the food?” and Speaker B replies “It’s certainly colorful”) was found to be the most difficult. Only 1 out of 54 lower-proficiency learners was able to comprehend this item correctly, and almost 80% of all learners were distracted by the second type of distractor. These results suggest that, when implied meaning is less accessible to learners, they tend to rely on their memory more in order to make choices.

In sum, this post hoc analysis demonstrated that the learners’ choice of distractors showed that L2 learners tended to be more attracted to the distractor that contained the words most recently heard. These findings help us understand where potential problems lie when learners make a wrong inference. When the target implied meaning is not salient, learners are confused by other irrelevant and misleading cues, particularly by the words that remain strongest in their short-term memory.

**Post Hoc Analysis of Effects of L1 Response Speed on L2 Response Speed**

In pragmatic comprehension, analyses of L2 proficiency impact on response time demonstrated no notable effect of general L2 proficiency on language processing speed. Thus, it was concluded that overall processing capacity for the pragmatic knowledge examined—namely the knowledge of how to interpret speaker intentions—did not reflect upon L2 proficiency. The question then remained as to what factors within the domain of ability for language use actually affected processing speed. These factors may well include cognitive and noncognitive factors separate from L2 knowledge. In order to address this question, an additional measure of L1 listening was administered to a subset of 59 L2 participants, on voluntary basis, in order to check whether the general speed of information processing in their L1 influenced the speed of processing in the L2. The test consisted of 12 items that were drawn from the Japanese Language Proficiency Test given annually to evaluate and certify Japanese proficiency of nonnative speakers seeking higher education in Japan (Association of International Education in Japan, 2003). The questions came from old versions of Level One, the highest grade test (years 1999 and 2000) and mainly assessed comprehension of literal meaning.

The format of the L1 listening test items mirrored that of the pragmatic listening items—a short conversation followed by a multiple choice question with four answer options. The L1 listening test was given through computers. Response time was measured from the moment the question appeared on the screen until the moment the participants pressed the number key for answer choice. Table 10 displays descriptive statistics of the response times in the two listening measures.

The average response time was more than two times faster for the L1 listening test, and the variance was also smaller, than for the L2 pragmatic listening task. A regression analysis was conducted in order to check whether L1 processing speed related to L2 processing speed. The results revealed no significant relationship between the two at the alpha level of .007; \(r = 0.34\). \(R^2\) of .12 also suggests that comparatively little variance in L2 response speed (i.e., 12%) was explained by L1 response speed.

**INTERPRETATIONS OF THE RESULTS**

This study inquired into the pragmatic L2 comprehension of Japanese adult learners of English. Differential comprehension loads according to the types of implied meaning, claimed in the previous literature, were explicitly tested in this

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pragmatic Listening Task</td>
<td>32</td>
<td>20.27</td>
<td>19.60</td>
<td>5.80</td>
<td>9.64</td>
</tr>
<tr>
<td>L1 Listening Test</td>
<td>12</td>
<td>9.29</td>
<td>8.56</td>
<td>3.64</td>
<td>4.00</td>
</tr>
</tbody>
</table>

**TABLE 10**

Response Time (in Seconds) for L2 Pragmatic Listening Task and L1 Listening Test \((N = 59)\)

Note. \(K\) = the number of items. Response time = average number of seconds taken to answer each item correctly.
research by analyzing comprehension of different item categories within one sample population. The present study was also motivated by previous findings that learners with higher L2 proficiency performed pragmatic functions better than learners with lower proficiency, and by the desire to assess comprehension rather than production of pragmatic functions. Pragmatic comprehension was operationalized as the ability to comprehend a speaker’s implied meaning, namely meaning that is not explicitly stated in an utterance.

**Interpretation of the Effect of Item Types on Comprehension**

The first research question addressed the influence of implied meaning types on comprehension in terms of accuracy and speed of comprehension. There was statistical evidence for L2 learners that comprehension of LCIs was more difficult and took longer than that of MCIs, whereas there was no such category difference for native speakers’ comprehension. The results suggest differential comprehension issues for these two types of implied meaning in EFL learning. Different degrees of transparency of meaning or accessibility to meaning exist in different implied meaning types, and differences are manifest in the contrast between L2 comprehension of MCIs and LCIs, as documented in this research. The differential comprehension load surfaced in accuracy, as well as in processing fluency when deriving accurate answers.

The findings of this study support claims about the role of conventionality in language comprehension. Use of conventional features in conversation speeds up normal communication by allowing the speaker to assume that the listener will enact normal implicatures. Familiarity with these conventions reduces the processing load on the listener as well, allowing the listener to spend processing resources on conversational information that is genuinely new. When features of conventionality are not used by the speaker, more processing effort is required by the listener. The listener then requires more time to draw inferences about the speaker’s intended meaning and the implications for the pragmatic context (i.e., how the speaker is representing information about the conversation topic).

In the present study, linguistic and nonlinguistic contexts encoded in MCIs may be more routinized and conventionalized than those in LCIs. The MCIs included relatively fixed patterns of discourse exchange and specific linguistic expressions or patterns used to realize the goals of speech acts. It has been claimed that these features of conventionality reduce processing effort in human communication, and this claim also proved true for L2 comprehension in this study. In contrast, the less conventionalized, idiosyncratic, and context-dependent nature of LCIs requires more extensive analytical processing of both linguistic and nonlinguistic cues in order to derive intended meaning. This claim was supported by the present findings because, even when accurately comprehended, the LCIs required significantly more time for learners to draw correct inferences. In the development of L2 pragmatic comprehension, then, the ability to comprehend MCIs might precede comprehension of LCIs due to the greater linguistic conventionality encoded in the former.

**Interpretation of the Effect of L2 Proficiency on Comprehension**

The second research question investigated the influence of general L2 proficiency on comprehension of implied meaning. Regression analysis revealed a significant L2 proficiency effect in accuracy of comprehension for both item categories with a relatively high determinant coefficient value. The findings suggest a distinct relationship between general L2 proficiency as measured by the institutional TOEFL and accurate comprehension of implicit, nonliteral utterances.

As shown in the post hoc error analysis, for L2 learners, the greatest source of distraction in comprehension seemed to lie in short-term memory. When the learners made a wrong choice, they were most attracted by the option containing words from the last utterance heard. This memory impact may be specific to listening tasks. Because the input in listening is acoustic and transient in nature, listening requires an automatic, real-time comprehension. Thus, listening could be greatly affected by learners’ short-term memory because, after listening, learners could not refer back to a text. Due to this memory constraint, listening comprehension occurs in “piece-meal” fashion (Foss & Hakes, 1978). When we listen to an utterance, we do not record every single word as it occurs. While we listen to an utterance, we transform its strings of words into some structural representation in the form of “chunks,” in order to retain information effectively as well as to free up the short-term memory to deal with new incoming information. The original utterance perceived is rapidly forgotten due to the demand on the short-term memory. Because higher L2 proficiency learners were more accurate in pragmatic comprehension, it is possible that, as
proficiency develops, learners may acquire a way to control short-term memory; they might become better able to direct their attention to the most relevant part of information and retain the gist of information in short-term memory, resulting in improved comprehension accuracy.

Although the results for comprehension accuracy support the findings of previous research that higher L2 proficiency leads to better pragmatic performance, proficiency has a mixed influence on the speed of comprehension. In the analyses of the current study, proficiency was found to be a significant predictor for comprehension speed of MCIs, but no significant proficiency impact was found on response time for LCIs. However, the overall effect size, indicated in the correlation coefficient value, was low (i.e., $R^2$-square of .11), suggesting that there was not much change in comprehension speed for the proficiency. Thus, this study did not directly support previous claims that comprehension speed is a consistent factor in L2 proficiency and development. Previous literature (e.g., Bialystok, 1990, 1994; Segalowitz, 2000) argued that automaticity in language processing is an emergent property of high-level language control. In this study, automaticity was evident in L1 comprehension because response speed was much faster and more uniform than for the learners’ L2. However, in this study, automaticity was not evident as a factor that discriminated among learners at the different proficiency levels examined. Previous research in the area of psycholinguistics has mainly focused on lower-order processing (e.g., word recognition, sentence decoding, shadowing) in making inferences about automatic and fluent control of language information. Thus, when higher-level processing, such as the decoding of indirect utterances, is combined with lower-level processing, performance speed might show a different developmental course, due to the greater number of linguistic, contextual, and sociocultural resources to be processed during performance.

Another possible explanation for the lack of proficiency impact on performance speed is the restriction of L2 proficiency range in this study, namely the absence of an advanced-level learner group. As Segalowitz et al. (1998) claimed, learners who are relative beginners differ among themselves primarily in what they know about language, rather than in their processing efficiency. It is only after processing sequences have become routinized that individual differences surface in processing fluency. This claim seems to apply to the present findings because the learners differed in accuracy scores, namely the knowledge of language, but not in their processing speed. Relatively greater levels of general L2 proficiency (i.e., above a threshold level) might be needed before any increase in speed is observed. The L2 learners in this study may not have achieved the threshold proficiency needed to exhibit differential processing speed.

Another potential reason for the lack of relationship between L2 proficiency and performance speed in this study is that speed may be strongly affected by factors other than proficiency. Individuals tend to differ in strategic competence (e.g., planning, goal setting, and decision-making) as well as in personality, affective, and cognitive factors (e.g., general information processing skills). Speed in pragmatic processing may also be influenced by a number of individual, nonlanguage factors other than mere linguistic processing, such as sociocultural knowledge, experience, and background, which eventually lead to greater individual variance. The much greater variance observed in the L2 time data than in the score data also supports the claim that individual variability influences processing time. Different individuals engage in different thought processes during performance and focus on different requirements of the task. They tend to differ in the ways in which they cope with processing demands in real time. One potentially influential factor, L1 processing speed, was examined in a post hoc analysis in the pragmatic listening task. No notable influence of L1 processing speed on L2 processing speed was found, which implies that general information processing speed in the L1 was not the factor that influenced L2 processing in the present sample.

When the processing speed difference between MCIs and LCIs was analyzed, although the magnitude of effect was low, proficiency was found to be a significant predictor for comprehension speed of MCIs, but not for LCIs. This finding could be due to the fact that the MCIs were simpler to understand because of their conventionality, as documented in the results for the first research question. Because the pragmatic processing for comprehending the MCIs was handled by higher proficiency L2 learners with relative ease, what was left for them in comprehension could be mainly linguistic processing (e.g., skimming through question and answer options), at which higher-level learners were more skilled and faster than lower-level learners. It is likely that, when difficulty of pragmatic decoding is relatively controlled, automaticity in overall language processing is promoted. As a result, proficiency might begin to make a difference in comprehension speed, because higher-proficiency learners tend to be more automatized at the lower-order
pragmatic skills, as documented in the previous literature. In contrast, processing speed for comprehending LCI items did not differ along with L2 proficiency because comprehension of pragmatic meaning required greater effort than the mere linguistic processing. Higher-level learners were more accurate, but took as much time as the lower-level learners to make a decision because there was a greater degree of ambiguity and uncertainty involved in comprehension.

**Interpretation of the Relationship between Accuracy and Speed**

The differential results between accuracy and comprehension speed were supported by the analysis of the third research question. Using partial correlation, the last question sought the relationship between accuracy and comprehension time of pragmatic meaning. When the effect of L2 proficiency was removed, no significant relationship was found between accuracy and processing speed for implied meaning. These results offer counterevidence to the accuracy-speed trade off, which was widely claimed in the field of cognitive psychology (Lachman et al., 1979; Oomen & Postma, 2001), as well as in L2 research (Skehan, 1998). Existing L1 studies generally confirmed that increasing speed increases errors, and increasing accuracy costs time. This study documented that such a negative relationship between accuracy and speed is not necessarily evident in L2 processing, particularly for the present L2 participant group with proficiency ranging from the true beginner to high-intermediate level. These results imply that fluent language processing deserves an independent analysis, apart from an analysis of accuracy, because accuracy and speed do not develop in a parallel manner in interlanguage pragmatics.

In summary, this study identified a strong proficiency effect on the accuracy but not on the speed of comprehension of pragmatic functions. The findings suggest that accuracy and processing speed may comprise two distinct dimensions of L2 competence in pragmatic comprehension. The ability to understand implied information and the ability to process the information quickly might exhibit different task characteristics and are influenced differently by L2 proficiency.

**IMPLICATIONS OF THIS STUDY FOR PRAGMATIC TEACHING**

This study offers some implications for explicit instruction of pragmatic skills, in terms of content and method of instruction. First, the study suggests that comprehension of more conventional indirect speech acts should be introduced and practiced prior to LCI items, because they were easier and faster to comprehend. Teachers can give examples of different types of implicatures and solicit discussion that helps learners focus on the intended interpretation. Teachers can also highlight conventional features associated with indirect speech acts, such as specific syntactic forms (e.g., *I am wondering if* + verb for a request) or linguistic patterns used to convey illocution (e.g., giving an excuse for a refusal), emphasizing that these features typically assist comprehension.

When teaching less conventional, idiosyncratic implicatures, however, the instructional focus should be placed on the overall analysis of both linguistic and contextual features in order to derive the intended meaning. Teachers should analyze contextual, grammatical, semantic, and sociolinguistic resources needed to comprehend implicatures and teach them to the students. For instance, teachers could use the concept of the adjacency pair rule, the basic structure of a conversation exchange, in order to explain that the second pair-part in an exchange is constrained by the first pair-part and provides relevant information even when the utterance presents a mismatch at the literal level. Teachers could also encourage learners to use contextual features, such as paralinguistic cues (e.g., intonation, tone of voice, pause length), background knowledge, and personal experience, in order to seek relevance in the seemingly irrelevant information. Such practices will exercise learners’ inferential skills beyond basic comprehension skills, and promote their awareness of indirect communication. A follow-up discussion on why speakers use indirect utterances instead of literal ones would also deepen learners’ understanding of the strategic intention behind implicatures. Flouting of the relevance maxim often appears as a strategic device for avoiding explicit answers in order to be polite and save face in social interaction. By understanding the logic behind such implicatures, learners will realize that implicatures are closely related to particular contexts and language users.

Another instructional implication relates to the fluency-accuracy balance in pedagogical contexts. Classroom teachers could benefit from the present findings that fluency and accuracy do not develop in a parallel manner, as observed in this sample of EFL learners. Thus, teachers should not expect the same rate of development in both dimensions nor should they set an overly idealistic
goal, such as “fluent accuracy,” as an immediate objective of teaching because fluency seems to lag behind accuracy in L2 development.

LIMITATIONS OF THE STUDY AND DIRECTIONS FOR FUTURE RESEARCH

Based on the limitations of the present study, several implications for future research are proposed. One major limitation of this study is that it did not include advanced-level learners in the subset of the L2 sample. In this study, the learners’ proficiency levels ranged from beginner to high-intermediate, with a TOEFL range between 330 and 590 and a mean of 457. Thus, adding an advanced learner group (i.e., TOEFL scores over 600) for the accuracy and speed analysis of pragmatic comprehension could prove interesting. It could be that advanced learners more clearly approximate native speaker performance, and as a result, the influence of proficiency might manifest more strongly in their processing speed, as well as in their accuracy of comprehension. It would be also interesting to compare learners in a foreign language environment versus learners in a L2 environment. It is possible that, in the L2 context, learners are more exposed to the everyday practice of comprehending implicit meaning, and thus their comprehension accuracy and speed might be better than those of EFL students with limited target language input.

Another limitation of this study is that the construct of pragmatic comprehension was limited to utterance-level comprehension. This study measured L2 inferential skills using a series of conversations that included utterance-level deviation from literal meaning. However, in real life conversations, implicitness or inexplicitness in meaning is manifest through various sequential and cumulative features of spoken discourse, such as deixis and reference words, ellipsis, incomplete sentences, or disfluency features (Cheng & Warren, 1999; Rost, 2002). Corpora of naturally occurring conversations should be analyzed closely in order to identify a set of conversational features that independently or jointly make up for overall explicitness or inexplicitness in conversation. In future research, those features could be incorporated into a listening task in order to measure L2 comprehension of pragmatic meaning.

This study is also limited because it sought a relationship between general L2 proficiency and comprehension ability of pragmatic functions, but it did not address developmental trends of pragmatic comprehension. Future research should incorporate longitudinal analyses in order to examine L2 learners of different proficiency levels over extended periods of time and to document a continuum of development in comprehension of implicatures.

Finally, this study found that L2 proficiency was not a factor that strongly influenced performance speed. Moreover, L1 processing speed was found not to be related strongly to L2 processing speed. A question then remains as to what factors actually affect processing speed in L2 pragmatic tasks. Some candidate factors belong to the domain of cognition, including short-term memory, decision-making style, and attention control (Bialystok, 1990; Robinson, 2003; Widdowson, 1989). These factors reflect overall cognitive control, as a set of components that enables us to allocate attention and to select, coordinate, and integrate information in real time. In addition to examining cognitive abilities, recent studies explored the influence of the learning environment (e.g., domestic vs. abroad) and the amount of language contact on the rapidity of L2 oral production (Freed, Segalowitz, & Dewey, 2004; Segalowitz & Freed, 2004). These longitudinal studies have shown that certain types of cognitive abilities, learning contexts, and L2 experiences are positively related to fluidity of oral production. Following this recent trend, future research that incorporates additional measures of cognitive factors and descriptions of learning experiences will reveal the precise factors that determine the fluency dimension of language performance in L2s. Such research will guide us toward a better understanding of the interdependence among language, cognition, and context, and consequently broaden inquiry into the nature of L2 pragmatic performance.

ACKNOWLEDGMENTS

This study is based on my dissertation, completed at Northern Arizona University in 2003. I am deeply indebted to my committee chair, Dr. Mary McGroarty, and also to committee members, Dr. Joan Jamieson, Dr. Susan Fitzmaurice, and Dr. Michelle Miller, for their guidance and sage advice throughout this study. Special thanks go to Mr. Thomas Blair and Dr. Don Nilson for their assistance with data collection, and also to Dr. Graydon Bell for his assistance with statistical analyses. I would also like to thank four anonymous MLJ reviewers for their helpful comments on earlier drafts. I am solely responsible for all the errors that may remain.
NOTES

1 An earlier version of this article was presented at the annual conference of the American Association for Applied Linguistics in Arlington, VA, in March 2003.

2 The filler items dealt with basic, literal comprehension, as shown in the simplified example below:

   Tom: How do you like Atlanta so far?
   Sally: I love it.
   Question: What does Sally think about Atlanta?
   (a) She thinks that Atlanta is a dirty city.
   (b) She hasn’t seen much of the city since she moved in.
   (c) She thinks the city needs more great changes.
   (d) She likes Atlanta and enjoys living there.

3 It was not possible to follow the three distractor principles for all items. For example, the first distractor principle was not applicable for the indirect request items because the distractor that is “opposite to the intention of request” appeared illogical and thus implausible as a distractor. Moreover, when the last utterance in the dialogue was extremely short, containing only a few words, it was difficult to write a distractor following the second principle, “taking words from the last utterance.”

4 Because the present task did not involve much computer skill, the participants did not show any difficulty in using the computer to complete the task.

5 The higher-proficiency group had a mean institutional TOEFL score of 503, ranging from 485 to 590 ($SD = 19.46$). The lower-proficiency group had a mean institutional TOEFL score of 398 ranging between 330 and 430 ($SD = 25$).

REFERENCES


Rost, M. (2002). Teaching and researching listening. New York: Pearson Education.


