Language Teaching (2026), 1–18
doi:10.1017/S0261444825101122

CAMBRIDGE UNIVERSITY PRESS

PLENARY SPEECH

Replication in L2 pragmatics

Kathleen Bardovi-Harlig

Department of Second Language Studies, Indiana University, Bloomington, IN, USA
Email: bardovi@indiana.edu

(Received 15 October 2025; accepted 29 December 2025)

Abstract

This paper examines replication research in pragmatics. The paper has three goals: to understand how replication has been used in pragmatics, to explore how replication research can enrich research in pragmatics and language learning, and to offer some suggestions for replication projects in L2 pragmatics. The paper examines sets of original and replicated studies in both L1 and L2 pragmatics to understand the range of research that has been conducted. It then considers the status of item replications (repeated scenarios) that characterize L2 pragmatics research. And it concludes by considering specific issues in L2 pragmatics research that can be insightfully investigated via replication.

Keywords: pragmatics; replication; research method

1. Introduction

This paper explores the use of replication as a research tool in second language pragmatics. Advocates of replication research point to advantages such as better understanding of the findings of previous studies, confirming results, and enhancing generalizability (Porte & McManus, 2019; McManus, 2024). Replication has also been credited with illuminating aspects of research design, which has been an ongoing concern in L2 pragmatics research since its conception (Bardovi-Harlig, 2025). Two levels of replication are relevant to the current discussion: close replications and approximate replications. In a close replication study, only one major variable is modified, and all other aspects of the initial study are kept as constant as possible. In an approximate replication, two variables may be modified (McManus, 2023). Porte and McManus (2019, p. 26) also caution researchers to assure that their replication study has 'the same or greater methodological and analytical rigor as that of the original'.

Earlier discussions of replication in applied linguistics and second language acquisition have observed that few replication studies have been published, although replication has become more frequent because of the efforts of some researchers to educate the field of its value (see e.g. Language Teaching Review Panel, 2008; McManus, 2022, 2024, inter alia; Polio & Gass, 1997; Porte, 2012; Porte & McManus, 2019; Valdman, 1993). This paper takes a different approach, investigating replication research in terms of what it can tell us about pragmatics. This paper has three goals: to understand how replication has been used in pragmatics, to explore how replication research can enrich research in pragmatics and language learning, and to offer some suggestions for getting started on a replication project in L2 pragmatics.

Replication research is not new in L2 pragmatics. Two replications were published in the early 2000s (Billmyer & Varghese, 2000; Niezgoda & Roever, 2001). Bardovi-Harlig (2010) used 'replicability' as one point of evaluation of 152 studies in L2 pragmatics in terms of whether studies provided

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike licence (http://creativecommons.org/licenses/by-nc-sa/4.0), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the same Creative Commons licence is used to distribute the re-used or adapted article and the original article is properly cited. The written permission of Cambridge University Press or the rights holder(s) must be obtained prior to any commercial use.

https://doi.org/10.1017/S0261444825101122 Published online by Cambridge University Press

sufficient information about method to allow subsequent replication. Of 111 studies that reported on experimental (vs observational) data, 72, or 65% provided sufficient information about method to support replication. Another indication of recent interest in replication in L2 pragmatics is found in Taguchi and Li's (2019) article which outlines how to replicate two studies researching contextual and individual influences in pragmatic competence (Bardovi-Harlig & Bastos, 2011; Taguchi et al., 2016b).

As we will see, replication studies in L2 pragmatics have been used to test task effects, interpret results, and investigate learning contexts. These studies have resulted in further improvement of elicitation tasks and other aspects of method, identifying task effects, confirming or extending results, and extending findings to other languages. The sections that follow examine replications of studies in pragmatics, discuss repetitions of specific scenarios across studies in light of replication, and conclude by offering suggestions for starting points for variable manipulation in L1 and L2 pragmatics studies.

2. The replications

Collecting replication studies raises the question of how we recognize replications. McManus (2022) suggests that they should be identified by 'replication' in the title (e.g. *Understanding indirect meaning: A close replication*, Knight & Edmonds, 2024) or in the abstract ('set out to *replicate*' Schauer, 2006). This search for replication studies in pragmatics started first with pragmatics journals (*Journal of Pragmatics* and *Intercultural Pragmatics*) and searched 'replicat- (-e, -ed, -ion, -ing)'. In applied linguistics journals, I searched 'replicat-' and 'pragmatics' anywhere in the article. This was followed by a search on Ebsco and Google Scholar for 'replicat-' and 'pragmatics'. Uses of 'replicated' to refer to results that were consistent with an earlier study rather than a replication of a study were excluded (see also Marsden et al., 2018, supplement 1, p. 15). All original studies were articles published in peer-reviewed journals. Nine self-reported replications were identified. One study was excluded because it did not replicate the instrument used in the original study.

It is worth noting that Knight and Edmonds (2024) were the only authors to identify their study as a replication in the title. Two studies were identified as replications in the abstract using 'replication' (Chen & Yang, 2010; Schauer, 2006). The book chapter (Niezgoda & Roever, 2001), lacking an abstract, identified the replication in the introduction. Idemaru et al. (2019) identified components of the replication, using the same method and comparing the studies, in the abstract (i.e. 'compared to previously published Korean data collected through the same methodology') and reported the study to be a replication in the introduction (p. 519). Licea-Haquet et al. (2019) also used a paraphrase in the abstract and 'replication' in the body of the paper. The other two articles used paraphrases, one referring to 'original' and 'extended' formats of items (Wiegmann, 2023), and the other to 'descriptions of situations […] formulated by [the original study]' (Billmyer & Varghese, 2000). Replications are probably under-reported (McManus, 2022) and so the studies discussed here are to be considered examples, but not an exhaustive account.[1] For that reason, I have included L1 empirical studies as well as L2 studies for breadth of coverage. (I will claim that replicating L1 studies by manipulating the L1/L2 variable would be informative; see Section 4.1.)

Eight replication studies from one edited volume and six journals, *Journal of Pragmatics, Intercultural Pragmatics, Foreign Language Annals, Analysis, Language Learning,* and *Applied Linguistics*, were identified and are discussed here. Table 1 provides an overview of the original studies and the replications. Two of the studies draw on two studies each for their replications. Idemaru et al. (2019) replicates procedures and tasks from Winter and Grawunder (2012) and Brown et al. (2014); Wiegmann (2023) replicates the items from Weissman and Terkourafi (2019) with the format used by Reins and Wiegmann (2021). Two replication studies, Niezgoda and Roever (2001) and Schauer (2006), replicate the same study, Bardovi-Harlig and Dörnyei (1998). Four of the replications are carried out by one or more of the original authors (indicated by underlining in Table 1).[2] The variables

https://doi.org/10.1017/S0261444825101122 Published online by Cambridge University Press

Table 1. Replication studies in pragmatics

Original studies	Replication Studies	Task	Variable
L1 studies			
Chen (1993)	Chen & Yang (2010)	Written DCT (multiple responses)	Diachronic variation (language change)
Holtgraves (2008)	Licea-Haquet, Velásquez-Upegui, Holtgraves, & Giordano (2019)	Speech act activation task	Language (English→Spanish)
a. Winter & Grawunder (2012); b. Brown, Winter, Idemaru & Grawunder (2014)	Idemaru, Winter, & Brown (2019)	Oral DCT, Script reading, perception (deference identification)	Language (Korean→Japanese)
a.Weissman & Terkourafi (2019); b. Reins & Wiegmann (2021)	Wiegmann (2023)	Interpretation task (written)	Same items; changed item format
L2 studies			
Rose (1992)	Billmyer & Varghese (2000)	Written DCT	Elaborated scenarios; added L2 learners
Bardovi-Harlig & Dörnyei (1998)	Niezgoda & Roever (2001)	Video pragmatic-grammaticality judgement task	Added separate explicit identification of pragmatic and grammatical items; retested host and home environments
Bardovi-Harlig & Dörnyei (1998)	Schauer (2006)	Video pragmatic-grammaticality judgement task	Interview regarding source of ratings; retested host and home environments
Taguchi, Gomez-Laich, & Arrufat-Marques (2016a)	Knight and Edmonds (2024)	Audio-visual interpretation of indirectness	Language (Spanish→ French)

Note: The underlined names indicate researchers who participated in both the original and replication studies.

that were manipulated by the replications include time (diachronic comparison), item format, subject population, language, and learning context. The descriptions of the studies do not capture all the details of the original or replicated designs but focus on the pragmatics of the studies.

2.1. Time: Diachronic comparison

Chen and Yang (2010) executed a close replication of Chen (1993) which studied compliment responses in Mandarin Chinese. They administered the same four-scenario written DCT that encouraged multiple alternative responses, changing no features of the task. The replication took place in Xi'an, in the same city and at the same university as the original study 17 years later. The variable that is manipulated is the year of data collection. The results show that there had been a noticeable shift in responses to compliments in Mandarin Chinese in the city of X'ian. Whereas Chen (1993) reported that the dominant response to a compliment in Mandarin was a rejection (95.7% of compliment responses), Chen and Yang (2010) found that the preferred response was acceptance (62.6%) followed by deflection (28.3%). The replication provides a picture of diachronic change in a relatively short time.

2.2. Refining method

Replications may also be undertaken for the purpose of studying the effect of task design. In the case of the two methodological investigations considered here, the variable that is manipulated is item format. Rose (1992) used a six-item written DCT to study the effects of including a hearer response.

https://doi.org/10.1017/S0261444825101122 Published online by Cambridge University Press

Half of the participants completed the DCT with no hearer response and the other half completed the DCT with a hearer response following a blank for a turn to be supplied by the participant. There was no significant difference in NS responses to the two DCT formats, but there was a slight tendency for longer requests and greater use of downgraders and supportive moves on the DCT with no hearer responses. Billmyer and Varghese (2000) conducted a partial replication of Rose (1992) using the DCT with no hearer responses. Rose (1992) used unelaborated scenarios to elicit requests as in Example (1). Billmyer and Varghese added a second DCT format, with elaborated scenario descriptions providing more information about settings and speakers, as in Example (2). They also manipulated the participant population by testing L2 English speakers as well as L1 speakers. Half of the L1 speakers and half of the L2 learners took Version 1 (Example 1), and half took Version 2 (Example 2).

1. Rose (1992), Notes (without hearer response)
   You missed class and need to borrow a friend's notes. What would you say?
   YOU:
2. Billmyer and Varghese (2000), elaborated Notes item
   You are at the end of a history class and you are sitting next to Tom Yates. You missed last week's class and need to borrow his notes. He has been in the same program as you for one year and you see him socially about once a month in a group. You will also be taking classes together in the future. He is a good note-taker and one of the best students in the class. You have borrowed his notes twice before for the same class and the last time you borrowed them he was reluctant to give them up. In two weeks you both have the final exam for your class. What would you say?

   YOU:
3. Sample learner responses
   a. Hi. I missed class. Would you mind if I handed your notes? (Short version, Notes, L2 speaker)
   b. Tom, would you mind if I want to borrow your notes? I know it's the third time that I asked for but I had no chance to join the class last week. I'll be happy to help you if you need my notes in other classes. (Long version, Notes, L2 speaker)

   (Billmyer & Varghese, 2000, p. 540)

The longer scenarios yielded longer responses, more supportive moves, and more personalized responses. An example like (3a) is more likely to occur after (1), and (3b) is more likely to occur after (2). The replication showed the importance of the amount of information provided in the scenario used to elicit the production.

A second exploration of item design comes from Wiegmann (2023) which replicated Weissman and Terkourafi (2019), using the items from Weissman and Terkourafi (2019) and the format from Reins and Wiegmann (2021) to investigate differences in findings. The studies investigated the question of whether false or deceptive implicatures are lies. Weissman and Terkourafi (2019) cite the following example from Meibauer (2005, p. 1380) to illustrate a false implicature.

4. The story of the mate and the captain

A captain and his mate have a long-term quarrel. The mate drinks more rum than is good for him, and the captain is determined not to tolerate this behaviour any longer. When the mate is drunk again, the captain writes into the logbook: Today, 11th October, the mate is drunk. When the mate reads this entry during his next watch, he is first getting angry, then, after a short moment of reflection, he writes

https://doi.org/10.1017/S0261444825101122 Published online by Cambridge University Press

into the logbook: Today, 14th October, the captain is not drunk.

The implication of the mate's entry is that it is unusual that the captain was not drunk (and therefore remarkable), although the statement itself is true: The captain was not drunk. Subsequent studies by Weissman and Terkourafi (2019) and Reins and Wiegmann (2021) experimentally investigated whether speakers regard false implicatures, like the one created by the mate's entry, to be lies. However, they found conflicting results using different formats. Weissman and Terkourafi (2019) had found that false implicatures are not necessarily lies, whereas Reins and Wiegmann (2021) had found that false implicatures are lies. In order to reconcile the differences, Wiegmann (2023) replicated Weissman and Terkourafi (2019) as Condition 1 (Original, Example 5) and ran Condition 2 (Explicit, Example 6) in which the same scenarios were extended to explicitly state the speaker's intention to deceive and the implicated content. The changes appear in square brackets in Example (6). The scenarios were followed by a question, 'Did the speaker lie?' and a seven-point Likert scale from 1 'definitely not a lie' to 7 'definitely a lie.'

5. Weissman and Terkourafi (2019), original format

(Hammer) Rumours have spread about an incident in the art studio yesterday. Alex was in the studio all day and saw Sarah, frustrated with a project, pick up a hammer, walk over to a statue and kick the statue over with her foot, causing it to smash all over the floor. The following day, Alex talks about the incident.

Mark: I heard Sarah had a meltdown in the art studio yesterday! What happened?

Alex: You should've been there! In a fit of rage, Sarah picked up a hammer and broke a statue.

6. Wiegmann (2023), explicit format

(Hammer) Rumours have spread about an incident in the art studio yesterday. Alex was in the studio all day and saw Sarah, frustrated with a project, pick up a hammer, walk over to a statue and kick the statue over with her foot, causing it to smash all over the floor. The following day, Alex talks about the incident [and wants Mark to believe that Sarah used the hammer to destroy the statue].

Mark: I heard Sarah had a meltdown in the art studio yesterday! What happened?

Alex: You should've been there! In a fit of rage, Sarah picked up a hammer and broke a statue.

[Mark comes to believe that Sarah used a hammer to destroy the statue.]

The replication showed that in judging the original scenarios, participants found that most cases were not examples of lying (as the original did), but that, in contrast, the participants judged the explicitly restated scenarios as lies (Wiegmann, 2023, p. 115). Thus, the replication was able to demonstrate that item format is responsible for different outcomes. Taken together, the replications in this section illustrate the importance of this design feature.

2.3. Replication to expand languages investigated

Most branches of linguistics are interested in exploring multiple languages to determine whether language characteristics are universal or restricted to particular languages, and pragmatics is no exception. Within L2 pragmatics, either the first languages (L1s) or the second languages (L2) may be the variable that is manipulated during the replication.

https://doi.org/10.1017/S0261444825101122 Published online by Cambridge University Press

1. Manipulating the L1

Both replications of Bardovi-Harlig and Dörnyei (1998) manipulated the L1 variable. Bardovi-Harlig and Dörnyei (1998) tested Hungarian L1 EFL learners in Hungary and mixed L1 ESL learners in the Midwest of the US. Niezgoda and Roever (2001) tested Czech L1 EFL learners in the Czech Republic and mixed L1 ESL learners in Hawaii. Schauer (2006) tested German L1 students at home (EFL) and abroad in the UK (ESL), thus creating a German L1 ESL group, in contrast to the mixed L1 ESL groups used in the two previous studies. Varying the L1 while keeping the L2 constant allows researchers to utilize the original task. (These studies are discussed in more detail in Section 2.4.)

2. Manipulating the focal or target language

Replications that manipulate the language variable may change the language under investigation. Manipulating the language of investigation requires that the original task be translated, checked for pragmatic appropriateness, and repiloted. In this section we consider three replications that manipulated the focal language and thus created new instruments using the original task as a model and following the procedures for task development detailed in the original studies.

Idemaru et al. (2019) investigated phonetic cues of deferential speech in Japanese by replicating two studies on Korean, asking whether the phonetics of Korean deferential speech are found in Japanese. The replication had both a production component (Winter & Grawunder, 2012) and a perception component (Brown et al., 2014).

Winter and Grawunder (2012) investigated what phonetic cues signal deferential speech using two oral production tasks: leaving a phone message to a high or low status addressee and an oral DCT with five paired scenarios (to a higher status and lower or equal status addressee). They found that Korean speakers lowered their pitch when speaking in the formal register. A second study, Brown et al. (2014), tested the perception of intended politeness levels through phonetic cues in the absence of morphological or lexical information. The stimuli were ten request situations addressed to a low and high addressee read by eight native speakers of Korean (8 x 10 × 2 levels), creating 160 items. One crucial utterance was extracted from each item for the perception task. Participants were asked to judge whether the speaker was speaking to someone 'above the speaker' or 'below the speaker'. They found that L1 Korean listeners could perceive politeness without morphological or lexical indicators.

In their replication study, Idemaru et al. (2019) also used an oral DCT to elicit oral production data, following Winter and Grawunder (2012). They expanded the original five paired scenarios to higher and equal status addressees (ten items) to six pairs (12 items). To prepare the perception stimuli, they translated the ten request scripts at two levels used by Brown et al. (2014) into Japanese, modifying four out of ten items slightly, adding one pair, and then having 20 native speakers record them (creating a larger production sample as well as serving as the stimuli for the perception study). Following Brown et al. (2014), Study 2 extracted the crucial utterances from the script-reading task used in Study 1. They selected recordings from eight speakers, matching the original 160 utterances. In the replication, participants heard paired sentences and indicated which one was spoken to a higher status addressee (a departure from the original study that asked participants to judge whether there was a lower or higher status addressee for each sentence).

Idemaru et al. (2019) compared a range of phonetic characteristics of deferential speech in Japanese to those previously reported for Korean. In Korean and Japanese deferential voice is quieter and breathier, and with less fluctuation in pitch and volume, 'it sounds calm and soft' (p. 517). The replication pointed to differences too. Korean deferential speech shows lower pitch, whereas Japanese showed no consistent relationship of deference and pitch. Speakers of Japanese correctly identified the level of deference in only 56%, but Korean listeners identified the levels more robustly, in 70% of the cases. The results of the replication also challenge previous claims that Japanese polite speech is high pitched, and consequently the findings also challenge the claim that deference (or polite speech) is universally associated with high pitch. The replication led to new information about deferential speech in general and about Japanese in particular.

https://doi.org/10.1017/S0261444825101122 Published online by Cambridge University Press

The second set of studies are psycholinguistic studies of the comprehension and processing of speech acts and the identification of the illocutionary force of an utterance in the absence of performatives. The original study was conducted in English (Holtgraves, 2008) and the replication in Spanish (Licea-Haquet et al., 2019). In Experiment 1, Holtgraves (2008) presented short conversations to L1 English speakers. In the experimental conversations, the last utterance performed a speech act such as thanking, 'I appreciate your help so much. I couldn't have done it without you'. In the control versions the last utterance was altered so that it did not perform the speech act but included most of the same words: 'He appreciated my help so much. He couldn't have done it without me'. In addition to 12 speech act conversations and 12 control conversations in which the probe word did not occur, 24 filler items were created in which the probe word did occur. There was a separate listening and reading condition. When participants were presented with a probe word that described the speech act, such as *Thank* in the example shown here, and were asked whether the word appeared in the sentence, participants were more likely to incorrectly report that the probe word had occurred in the speech act sentence. Moreover, they took significantly longer to reject the probe word after the speech act sentence. Holtgraves interpreted this as showing that comprehension entails speech act activation.[3]

Licea-Haquet et al. (2019) replicated the study with L1 Spanish speakers, using a translation of the written format of Experiment 1 in Holtgraves (2008), piloting the conversations as Holtgraves did for the original English version. In the first Spanish version, Licea-Haquet et al. used verbs in infinitive form as the probe word for the experimental (speech act) utterances and the controls, but in conjugated forms for the fillers. The native Spanish speakers did not show the same patterns as the English speakers, which Licea-Haquet et al. hypothesized could be due to the form of the probe words. When the task was revised, and all the probe words were presented as infinitives, Spanish speakers took significantly longer to answer the speech act trials than the controls. That is, it took respondents longer to rule out the speech-act naming probe when the speech act had been performed than when it had not, even though the word did not occur in the last utterance of the conversation, just as in the original study. However, there was no difference in accuracy between the experimental and control conditions. Thus, the replication corroborated the increase in processing time, but not the reduced accuracy in rejecting the non-occurring probe word. This replication also illustrates the challenges of translation, especially when languages use non-equivalent forms, and it raises the necessity of additional replications to test the impact of the morphological forms of the probes.

The third set of studies discussed here investigated the interpretation of indirect speech. As in the previous two sets of studies, the replication manipulated the language to be investigated. Both the original and the replication studies tested the interpretation of indirect refusals, indirect opinions, and ironic utterances, addressing the question, 'Do L2 [language] learners demonstrate different comprehension accuracy and response speeds across different types of indirect meaning?' Based on conventionality, the authors of the original study, Taguchi et al. (2016a), expected the order of difficulty to progress from indirect refusals to indirect opinions to irony. However, they found that for L2 learners of Spanish (95% of whom were L1 English speakers), indirect refusals and indirect opinions were of equal difficulty and ironic utterances were more difficult. The replicating study (Knight & Edmonds, 2024) tested this finding with L1 speakers of English learning L2 French with the additional goal of increasing the generalizability of the findings with another language.

Taguchi et al. (2016a) implemented a video task that included 48 items, including 12 items each of indirect refusals, indirect opinions, and ironic utterances, ten fillers which had literal interpretations, and two practice items. The video clips had four to five speaker turns. Each item on the video was followed by a multiple-choice question from which participants selected an interpretation. The video and the items were developed expressly for the study; the pre-pilot phase of development included 16 items for indirect refusals, indirect opinions, and ironic utterances, which were reduced to 12 each after two rounds of piloting with native speakers of Spanish.

https://doi.org/10.1017/S0261444825101122 Published online by Cambridge University Press

To develop a French equivalent of the task for their close replication of the original, Knight and Edmonds contacted the first author of the original study for the pre-pilot task with 16 items per indirect speech act. They then followed the procedures for developing the original task reported by Taguchi and colleagues. Knight and Edmonds (2024) translated the items, verified cultural appropriacy (e.g. changing 'pretzels' to 'ice cream' for a snack purchased at a theatre for the Movies item), piloted the items with L1 French speakers, revised problematic items, piloted them again, and filmed the new task using L1 speakers, developing the multiple-choice questions (with appropriate responses and alternatives) using the procedure outlined in the original study. The L2 learners of French scored the same on indirect refusals and indirect opinions which both scored higher than irony, as was the case with L2 learners of Spanish. This replication showed the same results as the original, regardless of language. Thus, it showed that the findings could be generalized to languages beyond Spanish.

All the replications in this section illustrate the value of expanding the languages that are investigated. Whether L1 or L2, this increases the generalizability of the claims. The replications by Idemaru et al. (2019), Licea-Haquet et al. (2019), and Knight and Edmonds (2024) which tested additional target or focal languages also emphasize the investment in detail and cross-cultural design required when a new language is investigated; the replications all re-piloted the translated tasks. The replicating authors retained the methodological rigor of the original studies while adapting the task to a new language and culture.

2.4. Manipulating learning context and testing interpretations

This section discusses two well-known replications of Bardovi-Harlig and Dörnyei (1998), namely Niezgoda and Roever (2001) and Schauer (2006). Bardovi-Harlig and Dörnyei tested the ability of L2 learners to recognize pragmatic violations through a video judgement test that tested pragmatic and grammatical awareness. The 20-item video task featured the interactions of two students in various encounters with other people during their school day. There were eight items that were pragmatically appropriate but ungrammatical, eight items that were grammatical but pragmatically inappropriate, and four items that were both grammatical and appropriate.

The video played every scenario twice. In the first presentation, participants were told to 'just watch the scene' and in the repeated presentation, they were told to 'watch and mark your answer sheet' (p. 244). An exclamation mark appeared on the screen before the last utterance which was to be judged. Participants were given a written answer sheet which contained only the number of the item and the final utterance. They indicated whether the sentence was good/appropriate or bad/inappropriate by checking a box, then decided how bad the sentence was on a Likert scale from 'not bad at all' to 'very bad' (p. 244). The task was administered to EFL learners in Hungary (L1 Hungarian) and mixed L1 ESL students in the Midwest of the US. The study reported that ESL learners recognized more pragmatic infelicities and rated them as being more serious than grammatical errors, and the EFL learners recognized more grammatical errors and rated them more as being more serious than pragmatic infelicities. Bardovi-Harlig and Dörnyei concluded that the learning context, ESL vs EFL, was the determining variable.

Niezgoda and Roever (2001) and Schauer (2006) both used the video task to elicit learner judgements. In addition to the original question, they each had an additional investigation of the original findings. The authors of the original study had assumed that learners would reject the pragmatic infelicities on the basis of pragmatics and the ungrammatical sentences on the basis of grammatical violations, but did not explicitly test the assumption. However, both replications tested the original assumption by implementing an additional task. I will discuss this aspect of the replication first. Both replications also manipulated the learning context (and with it the L1, discussed briefly in 2.3.1) to test whether learning context was the determining factor.

Niezgoda and Roever (2001) conducted a close replication, using the video task and the original worksheet. Niezgoda and Roever also expanded the study to investigate whether learners had

https://doi.org/10.1017/S0261444825101122 Published online by Cambridge University Press

responded to grammatical items based on their grammatical errors and pragmatic items based their pragmatic infelicities. After running the original task, Niezgoda and Roever then trained a subset of participants to identify grammatical and pragmatic errors. The participants completed the task a second time, classifying errors as either pragmatic or grammatical. Learners were able to distinguish the pragmatic and grammatical items on the original task, strengthening the interpretation of the original study. Schauer's (2006) replication additionally interviewed learners to determine why they had marked errors; learners were told how they had responded to each item and then were asked to explain why. Like Niezgoda and Roever (2001), Schauer also found that learners could distinguish pragmatic from grammatical errors. By implementing additional tasks, both replication studies provided evidence that allows the results of original task to be interpreted with confidence.

Both replications also manipulated the variable of learning context, testing the original conclusion that learning context was the causative variable. Manipulating the context variable also entailed changing the subject population (and hence the L1). Niezgoda and Roever tested Czech L1 EFL learners in the Czech Republic and mixed L1 ESL learners in Hawaii. The L1 Czech learners were the top 5% of all English-language learners in the Czech Republic and aspiring EFL teachers, and the ESL learners were described as being in Hawaii to learn English and enjoy the beaches. The results showed that at a high enough level of proficiency and/or motivation EFL learners could recognize pragmatic infelicities as well as ESL learners (at low enough proficiency). The results of the replication suggested that the difference in learning context alone is not sufficient to determine the outcome, in contrast to the original interpretation.

In an approximate replication, Schauer (2006) tested a different population, namely German L1 speakers studying in a university in Germany (EFL learners), and German L1 speakers abroad in the UK (ESL learners). This replication kept the L1 constant across learning contexts which eliminates an extraneous variable in both Bardovi-Harlig and Dörnyei (1998) and Niezgoda and Roever (2001). Schauer also tested the study-abroad students twice, nine months apart, which added a longitudinal component to the study. Schauer reported that, as in the original study, at the first test, the EFL learners were less aware of pragmatic infelicities than the ESL group was. Schauer calculated both an 'uncorrected' rate (used in the original) and a 'corrected rate' which included only the pragmatic infelicities and grammatical errors that the learners correctly identified in the interview. Both analyses supported the original analysis. Learners in the UK identified more pragmatic errors than grammatical errors, and the learners in Germany identified more grammatical errors. Although the corrected rates were lower than the uncorrected rates of the original study, the learners in the UK rated pragmatic infelicities as more serious than grammatical errors, and the learners in Germany rated grammatical errors more seriously, further supporting the findings of the original study. The higher proficiency of the German ESL learners at nine months is more directly comparable to the high-level learners in the first replication by Niezgoda and Roever (2001), but is also at odds with their findings.

The longitudinal component of the study showed that at nine months of the length of stay, the ESL learners in the UK identified more pragmatic and grammatical errors, approximating the native speakers' scores. The severity ratings for the pragmatic infelicities increased significantly by the end of the nine months, whereas the severity ratings for the ungrammatical items were not significantly different. Thus, the severity ratings of the learners in the ESL context were even greater at nine months than those of the learners in the EFL context, supporting the imbalance of the severity ratings found in the original study.

There is one change to the original method that is worth mentioning. Schauer altered the answer sheet, thereby possibly altering the task. Instead of providing only the utterance to be judged on the answer sheet, the revised answer sheet provided the target utterance *and* the immediately preceding sentence to stimulate recall of the scenarios during the interview that followed. The original study provided *only* the last utterance to compel participants to watch the video in order to complete the task, judging the utterance in the video context rather than imagining the speakers, their prosody,

https://doi.org/10.1017/S0261444825101122 Published online by Cambridge University Press

or other relevant characteristics of the exchange as people may do in a written task. Providing the preceding utterance presents an adjacency pair (an utterance and the reply) on the answer sheet and this may have been sufficient to allow learners to exclusively focus on the written text if they were inclined to do so. The ramifications of changing the task are not discussed in the report. Given the advanced level of the learners, it is likely that the German L1 learners in Schauer's replication study did watch the video, but this may not be the case in all learning contexts or at all proficiency levels. This could be investigated in a subsequent study.

The Niezgoda and Roever (2001) and Schauer (2006) replications confirmed that the respondents were evaluating pragmatic infelicity and ungrammaticality as intended. Schauer also demonstrated that typical ESL learners showed greater pragmatic awareness than EFL learners, confirming the original findings. Niezgoda and Roever showed that exceptional EFL learners were not constrained by context, challenging the original interpretation. Both replications add to our understanding of the role of learning contexts, an area that remains of interest in L2 pragmatics research.

This review of the eight replication studies shows the range of replications that have been undertaken in pragmatics. The studies illustrate how replications can be used to test (and confirm) findings and investigate method and item construction, diachronic change, L1 and learning contexts, and different target languages. They provide examples for how additional replications may be carried out in L2 pragmatics. This set of replication-original pairs and triples also show that replications can lead to additional questions that require further empirical investigation.

3. Item replications

In addition to the replications at the level of a study, there are also item or scenario replications throughout the L2 pragmatics literature. This is possible because L2 pragmatics research is dominated by scenario-based elicited-production tasks (e.g. see Examples 1–2 and 7–10). While the predominant task is still the written DCT, oral DCTs and role plays are also scenario-based. Combined, they make up 172 of 246 (or 70%) of all the tasks and 172 of the 217 (or 79%) production tasks in Nguyen's (2019) methodological review. Item replications are not overtly described as replications and may not be thought of as replications by the researchers. The results from the individual items are not compared to the results for the same items in other studies. Yet both earlier and later studies use the scenarios to answer the question of how the particular speech act is realized, and addressing the same question is the most basic element of replication. These repeated items weave pragmatics studies together in a way not found in other areas of SLA research.

The repeated scenarios often draw on the some of the same scenarios that were introduced by the cross-cultural speech act realization pattern (CCSARP; Blum-Kulka et al., 1989; Blum-Kulka & Olshtain, 1984) project. In L2 pragmatics, the scenarios are often given short names so that they can be easily referred to in a paper; this means that they can also be identified across different studies. I will consider two scenarios here, 'Notes' and 'Forgot Book' (or 'Book'). The Notes scenario in Example (7) was used in the CCSARP and predates Examples (1) and (2) in Section 2.2. Notice that the CCSARP item format used third person and a rejoinder (a response from the hearer). The practice of using third-person wording in scenarios had changed to using second person by 1990 (e.g. Beebe et al., 1990; Hudson et al., 1992; Rose, 1992).

7. Notes (CCSARP, S5, Request to borrow notes)

At the university
Ann missed a lecture yesterday and would like to borrow Judith's notes.
Ann:_____
Judith: Sure, but let me have them back before the lecture next week.
(Blum-Kulka et al., 1989, p. 14)

https://doi.org/10.1017/S0261444825101122 Published online by Cambridge University Press

Table 2. Use of 'Notes' in L2 pragmatics studies by speech act, task, and target language

Study	Speech Act	Target Language	L1	Task
Johnston et al. (1998)	Request	English	Chinese	Written DCT
Billmyer & Varghese (2000)	Request	English	Mixed (7 languages)	Written DCT
Omar (2006)	Request	Kiswahili	English*	Written DCT
Félix-Brasdefer (2007)	Request	Spanish	English	Oral role-plays
Martínez-Flor (2008)	Request	English	Spanish*	Oral role-plays
Bella (2012)	Request	Greek	Mixed (14 languages)	Written judgement task
Su & Ren (2017)	Request	Chinese	English	Oral role-play
Beebe et al. (1990)	Refusal	English	Japanese	Written DCT
Félix-Brasdefer (2004)	Refusal	Spanish	English	Oral role-plays
Keshavarz et al. (2006)	Refusal	English	Persian	Written DCT
Chang (2011)	Refusal	English	Chinese	Oral DCT
Safont-Jordà & Portolés-Falomir (2013)	Refusal	English	Spanish and Catalan/Spanish	Oral role-plays

Note: * Indicates that students were from an American and a Spanish university, respectively, but L1 information was not provided.

The Notes scenario has been used to elicit both requests and refusals (Table 2). For requests, the scenarios may portray the requester sympathetically, as someone whose request for the notes is legitimate. In contrast, to elicit refusals the requester may be portrayed less sympathetically to assure that the speaker will refuse the request. Example (8) provides an elaborated Notes scenario used by Félix-Brasdefer (2004) to elicit refusals in Spanish (see also Example 2).

8. Elaborated Notes scenario used to elicit refusals in Spanish (Félix-Brasdefer, 2004)

   Imagine that you are in (Spanish-speaking country of your preference). You are taking a course in Latin American literature this semester. You haven't missed this class once this semester and consider yourself a diligent student. So far you have a good average in the class, not because it is easy for you, but because you have worked very hard. Among your classmates, you have a reputation for taking very good notes. The professor has just announced that the mid-term exam is next week.
   One of your classmates, who is taking a class with you for the first time this semester and who has frequently missed the class, asks you for your notes. You haven't interacted with him outside the class, but have occasionally done small group work together in class. When the class ends, he approaches you for your notes, but you don't want to lend them to him.

A second illustration of repeated scenarios in the L2 pragmatics literature is the 'Forgot Book' scenario used to elicit apologies (Table 3). In different studies the situation is described as forgetting an instructor's book or a peer's book. The original from the CCSARP which has written conversational turns is presented in Example (9); other items present only a scenario as in Example (10).

https://doi.org/10.1017/S0261444825101122 Published online by Cambridge University Press

Table 3. The use of 'Forgot book' in studies of L2 apologies

Study	Book owner	Target Language	L1	Task
Cohen & Olshtain (1993)	Classmate	English	Hebrew and Hebrew/French	Oral role-play
Bardovi-Harlig & Dörnyei (1998)	Classmate	English	Hungarian (EFL); mixed ESL	Audio-visual judgement task
Sabaté i Dalmau and Curell i Gotor (2007)	University teacher/ professor	English	Catalan	Written DCT
Warga & Schölmberger (2007)	Professor	Austrian German	French	Written DCT
Beckwith & Dewaele (2008)	Lecturer	Japanese	English	Written DCT
Bardovi-Harlig (2009)	Friend	English	Mixed	Oral DCT
Al Masaeed et al. (2018)	Professor	Arabic	English	Oral DCT
Barron (2019)	Professor	German	English	Written DCT
Chang & Su (2025)	Roommate	Mandarin (Beijing and Taiwan)	English	Oral DCT

9. Forgot book (CCSARP, S4, Apology for forgetting a borrowed book)

   At the college teacher's office
   Teacher: I hope you brought the book I lent you
   Miriam:
   Teacher: OK, but please remember it next week.
   (Blum-Kulka & Olshtain, 1984, p. 198; Blum-Kulka et al., 1989, p. 14)

10. Forgot book (Beckwith & Dewaele, 2008, p. 21)

    Imagine you are a student. You borrowed a book from your university lecturer, but you forgot to return it. What do you say to your lecturer?

Some articles do not provide the full scenario that was given to the particpants, but instead provide a summary such as 'a student forgets a professor's book' (e.g. Al Masaeed et al., 2018; Sabaté i Dalmau & Curell i Gotor, 2007). The repeated scenarios have been used in a variety of tasks, including written DCTs, oral DCTs, oral role-plays, and less frequently, judgement tasks (Tables 2 and 3). Articles do not typically cite other studies that have used the same basic scenario; however, attributing the sources of scenarios would enhance methodological transparency. See, for example, documentation of sources of scenarios for eliciting conventional expressions in Bardovi-Harlig and Su (2018) and Bardovi-Harlig et al. (2025). Providing the full task with all the scenarios as presented to the participants in publications may also encourage and facilitate replications.

In the aggregate, these repeated items could illustrate similarities and differences in L2 acquisition, cross-linguistic use, and different L1–L2 pairs. More to the point in terms of replication, understanding the repeated scenarios as replications would allow the aggregated studies to confirm findings across studies and test the generalizability of the findings of individual studies, two goals of replication. The repeated use of scenarios also shows a certain openness to replication in L2 pragmatics, although the item replications have some way to go to be real replications, even on their small scale.

https://doi.org/10.1017/S0261444825101122 Published online by Cambridge University Press

4. Routes to replication in L2 pragmatics

The replication studies in Section 2 provide examples of how replications of pragmatics studies have been designed and implemented with respect to the initial studies. Porte and McManus (2019) offer 18 characteristics that might make a study a candidate for replication (pp. 23–24). Researchers are advised to consider the relevance of a study, access to the original data collection materials, unexpected or unusual outcomes, familiarity of the method (to the researcher), and the original publication venue. Perhaps most important is their admonition to researchers to 'select your target studies for replication with an eye to helping move forward your own research practice as well as giving service to current knowledge in the field' (Porte & McManus, 2019, p. 15).

This section focuses on variables that might influence the pragmatic outcomes (other variables such as statistical analyses are discussed by Porte & McManus, 2019; McManus, 2022, 2023, *inter alia*). Drawing on the review of replications in Section 2, there are five particularly promising variables that can be manipulated in the replication of pragmatics studies: speaker population (testing L2 learners using tasks that have exclusively tested native speakers), the first language of L2 learners, the learning contexts, the modality of tasks, and target language. Each is considered in turn.

4.1. Manipulating the subject population by testing L2 learners

Empirical studies in L1 pragmatics may offer new perspectives for L2 studies. In his 1992 study, Rose tested the effects of item design exclusively on L1 speakers, whereas Billmyer and Varghese (2000) included L2 speakers as well as L1 speakers showing that the change in the scenarios affected both groups. There are also L1 studies that investigate areas of pragmatics that L2 studies have not yet investigated. For example, replications of the studies reported here on interpreting false implicatures (Reins & Wiegmann, 2021; Weissman & Terkourafi, 2019; Wiegmann, 2023) would expand the existing L2 research on implicature (and lies, which have not been investigated in L2 pragmatics).

Although the explicit identification of speech acts has been investigated with L2 learners (e.g. Koike, 1989, 1996), speech act identification during comprehension has not. Holtgraves (2008) and Licea-Haquet et al. (2019) have shown that speech act identification is activated during comprehension in native speakers in both English (Holtgraves, 2008) and Spanish (Licea-Haquet et al., 2019), but this has not yet been tested for L2 learners using the speech act probe described in Section 2.3.2.[4] Moreover, their design could additionally provide a novel means of exploring the comprehension of direct and indirect speech acts which were not a target of investigation in the original studies.

4.2. Expanding the first language of L2 learners

Expanding the range of L1s represented in L2 pragmatics research is a natural area for investigation. Investigating the acquisition of any single target language begs for a variety of L1 speakers to determine what is specific to particular L1–L2 pairs and what is universal. SLA research and L2 pragmatics research have both been interested in L1 influence, and expanding the range of L1 participants that complete the same task is one way to test it. If the original task was conducted in the target language, delivering it to another learner population requires no modification of the task (see, e.g. Bardovi-Harlig & Dörnyei, 1998, and its replications; Niezgoda & Roever, 2001; Schauer, 2006).

4.3. Manipulating the learning context

Changing the learning context of a study for a replication may also involve changing the first language(s) of the participants, as discussed in the previous section. Nevertheless, L1 and learning context are variables that can also be considered separately. Niezgoda and Roever (2001) and Schauer (2006) are good examples of studies that changed the learning contexts. Niezgoda and Roever's learners were Czech EFL learners and mixed L1 ESL students, but in different programs. In contrast,

https://doi.org/10.1017/S0261444825101122 Published online by Cambridge University Press

Schauer's study manipulated only the context, keeping the L1 constant. The learners were German L1 learners who were in the UK and at home in Germany.

Learning contexts in SLA studies usually refer to host environments or foreign language learners studying at their home institution. But given the dominance of university students in our studies (an issue in SLA more generally), we could include non-degree seeking participants. For example, Bella (2012) collected data using a written DCT modeled on Bardovi-Harlig and Dörnyei (1998) from economic migrants learning Greek in Greece. Studies of L2 pragmatics in New Zealand have included newcomers heading to the workforce (Riddiford & Joe, 2010). Expanding the range of L2 learning contexts and participants not only tests the generalizability of study results, but also strengthens SLA inquiry.

4.4. Manipulating the spoken/written variable for conversational pragmatics

Given my advocacy for oral tasks in L2 pragmatics research (Bardovi-Harlig, 2025; Bardovi-Harlig & Hartford, 2005, inter alia), it will come as no surprise that I think a promising line of replication research in L2 pragmatics would involve manipulating the variable of task modality, specifically in the direction of matching modality. Matching modality (Bardovi-Harlig, 2018) is a simple concept that refers to matching the modality of research instruments or tasks to the modality of the language event being investigated. In the study of conversational pragmatics, this would mean that tasks would be oral/aural. In addition, there is pressure to use oral tasks from investigations of pragmaprosody which emphasize the role of speaking in pragmatics (e.g. Hao et al., 2024; Kang et al., 2021; Liu, 2025). Many comprehension and interpretation studies ask participants to read written conversations, and many production studies ask participants to write conversational turns. Manipulating the variable of modality, studies that asked participants to read utterances or conversations and then judge pragmatic appropriateness, determine the illocutionary force or implicature, or interpret meaning would be presented aurally and participants would hear the discourse to be judged. Similarly, production tasks that ask participants to write their responses in the original study would ask them to speak in the replication. Additionally, other areas of SLA recognize that untimed written tasks are more likely to tap explicit (declarative) knowledge, whereas timed oral tasks are likely to draw on implicit (procedural) knowledge. Matching modality would address this concern as well, particularly in the oral/aural condition while increasing ecological validity of tasks that represent conversation.

4.5. Expanding language coverage by exploring other target languages

Expanding the language of investigation is perhaps both the most obvious and the most challenging variable for manipulation. It is obvious, yet worthwhile, to investigate whether what we have discovered about the acquisition of L2 pragmatics based on one language holds for other languages. This is particularly important given that some languages have been researched more than others. As we have seen from the examples set by Knight and Edmonds (2024), Idemaru et al. (2019), and Licea-Haquet et al. (2019), producing faithful reproductions of the original task in another language requires not only a translation of the task, but cultural adjustments, and then piloting, revising, recording, and sometimes filming the task in the new target language.

The range of possible cultural adaptations is indeed large, but in terms of a single item or task, they could be small but meaningful. Attention to detail is key. Considering only the authors already cited, I will give four examples. Knight and Edmonds (2024) changed the snack in their Movies scenario to a snack available in movie theatres in France (i.e. ice cream). Bella (2012) describes a request scenario using her knowledge of the local Greek public transportation system. The next two adaptations were made in studying regional pragmatic variation rather than in replications but are also relevant here. When studying regional pragmatic variation in Spanish, Bardovi-Harlig et al. (2025) changed the desired purchase from a blanket in El Paso/Ciudad Juarez to sunglasses in Barcelona in a bargaining

https://doi.org/10.1017/S0261444825101122 Published online by Cambridge University Press

scenario at an outdoor market. Su and Chang (2019) presented the apology scenarios in an oral DCT to Mandarin speakers in Mainland China in simplified Chinese characters and voiced by a speaker of standard Mandarin and they presented the same oral DCT to Mandarin speakers in Taiwan in traditional Chinese characters and voiced by a speaker of Taiwanese Mandarin. Each of these cultural adaptations make the scenarios they describe more familiar and thus more believable to the intended participant group when using tasks to elicit data from speakers and learners.

Developing interpretation or comprehension tasks requires a significant effort because they involve developing scenarios and capturing authentic or near-authentic language. In contrast, production tasks such as oral DCTs or role-plays require only the translation and cultural adjustments in the scenarios, and whatever initiating oral turns may be included in the task. The preparation of prompts is therefore less extensive than preparing a judgement, interpretation, or comprehension task. It is worth noting that replicating a study that used observational (production) data requires no translation. Although challenging, manipulating the target-language variable is necessary for understanding the limits of generalizability that may result from studying a small range of target languages.

4.6. Other considerations

In addition to Porte and McManus's (2019) admonition to maintain methodological rigor in the replicating studies, other considerations arise. Glaser (2020) raises the issue of updating scenarios used in the original studies. As she correctly points out, the scenario used by Bardovi-Harlig and Dörnyei (1998) in which a teacher asks a student 'to stop by the bus stop on your way home' to check the bus schedules for the class trip has been rendered obsolete by the internet. Bus schedules are easily found on websites. Glaser modifies her analogous item to read 'Lisa, could you do an online search and check out the …'? (p. 64). (Perhaps even this modification has been replaced by apps.) A similar item was unexpectedly discovered in Bardovi-Harlig (2009). The oral DCT modelled an item on a multiple-choice item from Roever's (2005) routines test. The scenario, leaving a message with a roommate, was designed to elicit the expression 'Can I leave a message?' However, by the time of data collection in 2006, the item was no longer relevant. Speakers no longer leave messages at a central location at friend's homes or businesses; voicemail has replaced this speech act. These examples show that changes in technology – like cultural differences – may need to be considered for replications.

While Glaser's (2020) observation highlights potential changes to specific items in a task, the results reported in Chen and Yang's (2010) study raise the issue of pragmatic diachronic variation in the language produced. We know that language changes, but we might not be fully aware of how quickly pragmatics may change (after all, we have only recently embraced synchronic variation in pragmatics, e.g. Barron & Schneider, 2009). Chen and Yang (2010) revealed a change in preferred responses to compliments at the level of pragmatic strategy. This suggests that it might be worthwhile to replicate some of our classic studies in L2 pragmatics to explore language changes and concomitant changes in language produced by L2 learners.

These are only a few variables that can be explored via close or approximate replications of L2 pragmatic studies. Taguchi and Li (2019), discussing potential replications of Taguchi et al. (2016b) and Bardovi-Harlig and Bastos (2011), additionally suggest varying proficiency, widening the range of length of stay, and individual differences whether cognitive (such as aptitude and working memory) or social-affective (such as motivation, willingness-to-communicate, and personality). Naturally, the ultimate identification of variables to be manipulated depends on the original study and the questions that lead to the replication.

5. Conclusion

By examining replication studies in L2 pragmatics, I hope to have raised awareness of the value of replication research among L2 pragmatics researchers and to have raised awareness of L2 pragmatics and its potential value as a site for replication among researchers already interested replication

https://doi.org/10.1017/S0261444825101122 Published online by Cambridge University Press

research. I have provided some ideas for variables that might be manipulated in L2 pragmatics replication studies. The replication studies reviewed here continue the focus on task design, which has been a crucial concern to L2 pragmatics research both in original and replication studies.

Notes

1. Based on responses by applied linguists to a survey about their replication practices, McManus (2022) estimates that as many as 50% of replications may not be labelled as replications.
2. McManus (2022) also found that just over half of the applied linguists who had experience with replication reported having replicated one or more of their own studies.
3. Holtgraves (2008) repeated the study using the same items with a lexical decision task in Experiment 2 of the same paper with comparable results.
4. Holtgraves (2007) tested a small group of L2 learners employing the lexical decision task used in Holtgraves (2008), Experiment 2. Learners could identify speech acts but not during online processing. This also seems worthy of replication with high proficiency L2 learners or a cross-sectional design.

References

Al Masaeed, K., Waugh, L. R., & Burns, K. E. (2018). The development of interlanguage pragmatics in L2 Arabic: The production of apology strategies. *System*, *74*, 98–108. https://doi.org/10.1016/j.system.2018.03.001

Bardovi-Harlig, K. (2009). Conventional expressions as a pragmalinguistic resource: Recognition and production of conventional expressions in L2 pragmatics. *Language Learning*, *59*(4), 755–795. https://doi.org/10.1111/j.1467-9922.2009.00525.x

Bardovi-Harlig, K. (2010). Exploring the pragmatics of interlanguage pragmatics: Definition by design. In A. Trosborg (Ed.), *Pragmatics across languages and cultures* (Vol. 7 of *Handbooks of pragmatics*, pp. 219–259). Mouton de Gruyter.

Bardovi-Harlig K. (2018). Matching modality in L2 pragmatics research design. System, 75 13–22. 10.1016/j.system.2018.03.007

Bardovi-Harlig, K. (2025). Research methods in L2 pragmatics research: Introduction to the special issue. *Research Methods in Applied Linguistics*, *4*(3), 1–8.100241. https://doi.org/10.1016/j.rmal.2025.100241

Bardovi-Harlig, K., & Bastos, M. T. (2011). Proficiency, length of stay, and intensity of interaction and the acquisition of conventional expressions in L2 pragmatics. *Intercultural Pragmatics*, *8*(3), 347–384. https://doi.org/10.1515/IPRG.2011.017

Bardovi-Harlig, K., Comajoan-Colomé, L., Mossman, S., & Rodríguez Sánchez, E. (2025). Regional pragmatic variation in the use of conventional expressions at three Spanish-speaking sites. *Lingua*, *326*, 1–25, 104021. https://doi.org/10.1016/j.lingua.2025.104021

Bardovi-Harlig, K., & Dörnyei, Z. (1998). Do language learners recognize pragmatic violations? Pragmatic vs. grammatical awareness in instructed L2 learning. *TESOL Quarterly*, *32*(2), 233–262. https://doi.org/10.2307/3587583

Bardovi-Harlig, K., & Hartford, B. S. (2005). Institutional discourse and interlanguage pragmatics. In K. Bardovi-Harlig, & B. S. Hartford (Eds.), *Interlanguage pragmatics: Exploring institutional talk* (pp. 7–36). Erlbaum.

Bardovi-Harlig, K., & Su, Y. (2018). The acquisition of conventional expressions as a pragmalinguistic resource in Chinese as a foreign language. *Modern Language Journal*, *102*(4), 732–757. https://doi.org/10.1111/modl.12517

Barron, A. (2019). Using corpus-linguistic methods to track longitudinal development: Routine apologies in the study abroad context. *Journal of Pragmatics*, *146*, 87–105. https://doi.org/10.1016/j.pragma.2018.08.015

Barron, A., & Schneider, K. P. (2009). Variational pragmatics: Studying the impact of social factors on language use in interaction. *Intercultural Pragmatics*, *6*(4), 425–442. https://doi.org/10.1515/IPRG.2009.023

Beckwith, S., & Dewaele, J. M. (2008). The development of apologies in the Japanese L2 of adult English native speakers. *Birkbeck Studies in Applied Linguistics*, *3*, 1–26. https://eprints.bbk.ac.uk/id/eprint/2683/

Beebe, L. M., Takahashi, T., & Uliss-Weltz, R. (1990). Pragmatic transfer in ESL refusals. In R. Scarcella, E. Anderson, & S. Krashen (Eds.), *Developing communicative competence in a second language* (pp. 55–73). Newbury House.

Bella, S. (2012). Pragmatic awareness in a second language setting: The case of L2 learners of Greek. *Multilingua*, *31*(1), 1–33. https://doi.org/10.1515/mult.2012.001

Billmyer, K., & Varghese, M. (2000). Investigating instrument-based pragmatic variability: Effects of enhancing discourse completion tests. *Applied Linguistics*, *21*(4), 517–552. https://doi.org/10.1093/applin/21.4.517

Blum-Kulka, S., House, J., & Kasper, G. (1989). *Cross cultural pragmatics: Requests and apologies*. Ablex.

Blum-Kulka, S., & Olshtain, E. (1984). Requests and apologies: A cross-cultural study of speech act realization patterns (CCSARP). *Applied Linguistics*, *5*(3), 196–213. https://doi.org/10.1093/applin/5.3.196

Brown, L., Winter, B., Idemaru, K., & Grawunder, S. (2014). Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues. *Journal of Pragmatics*, *66*, 45–60. https://doi.org/10.1016/j.pragma.2014.02.011

Chang, Y., & Su, Y. (2025, May 8–10). The pragmatic development of L2 apology in Mandarin Chinese. Paper presented at Chinese Language Teachers Association Annual Conference, New York City, New York.

https://doi.org/10.1017/S0261444825101122 Published online by Cambridge University Press

Chang, Y. F. (2011). Refusing in a foreign language: An investigation of problems encountered by Chinese learners of English. *Multilingua*, *30*(1), 71–98. https://doi.org/10.1515/mult.2011.004

Chen, R. (1993). Responding to compliments: A contrastive study of politeness strategies between American English and Chinese speakers. *Journal of Pragmatics*, *20*(1), 49–75. https://doi.org/10.1016/0378-2166(93)90106-y

Chen, R., & Yang, D. (2010). Responding to compliments in Chinese: Has it changed? *Journal of Pragmatics*, *42*(7), 1951–1963. https://doi.org/10.1016/j.pragma.2009.12.006

Cohen, A. D., & Olshtain, E. (1993). The production of speech acts by EFL learners. *TESOL Quarterly*, *27*(1), 33–56. https://doi.org/10.2307/3586950

Félix-Brasdefer, J. C. (2004). Interlanguage refusals: Linguistic politeness and length of residence in the target community. *Language Learning*, *54*(4), 587–653. https://doi.org/10.1111/j.1467-9922.2004.00281.x

Félix-Brasdefer, J. C. (2007). Pragmatic development in the Spanish as a FL classroom: A cross-sectional study of learner requests. *Intercultural Pragmatics*, *4*(2), 253–286. https://doi.org/10.1515/IP.2007.013

Glaser, K. (2020). Assessing the L2 pragmatic awareness of non-native EFL teacher candidates: Is spotting a problem enough? *Lodz Papers in Pragmatics*, *16*(1), 33–65. https://doi.org/10.1515/lpp-2020-0003

Hao, Y. C., Su, Y., & Chang, Y. (2024). Perceptual and acoustic analysis of prosody in Mandarin Chinese refusals. *Journal of Pragmatics*, *233*, 3–20. https://doi.org/10.1016/j.pragma.2024.09.004

Holtgraves, T. (2007). Second language learners and speech act comprehension. *Language Learning*, *57*(4), 595–610. https://doi.org/10.1111/j.1467-9922.2007.00429.x

Holtgraves, T. (2008). Automatic intention recognition in conversation processing. *Journal of Memory and Language*, *58*(3), 627–645. https://doi.org/10.1016/j.jml.2007.06.001

Hudson, T., Detmer, E., & Brown, J. D. (1992). *A framework for testing cross-cultural pragmatics*. University of Hawai'i at Manoa, Second Language Teaching and Curriculum Center.

Idemaru, K., Winter, B., & Brown, L. (2019). Cross-cultural multimodal politeness: The phonetics of Japanese deferential speech in comparison to Korean. *Intercultural Pragmatics*, *16*(5), 517–555. https://doi.org/10.1515/ip-2019-0027

Johnston, B., Kasper, G., & Ross, S. (1998). Effect of rejoinders in production questionnaires. *Applied Linguistics*, *19*(2), 157–182. https://doi.org/10.1093/applin/19.2.157

Kang, O., Kermad, A., & Taguchi, N. (2021). The interplay of proficiency and study abroad experience on the prosody of L2 speech acts. *Journal of Second Language Pronunciation*, *7*(3), 343–369. https://doi.org/10.1075/jslp.20024.kan

Keshavarz, M. H., Eslami, Z. R., & Ghahraman, V. (2006). Pragmatic transfer and Iranian EFL refusals: A cross-cultural perspective of Persian and English. *Pragmatics and Language Learning*, *11*, 359–402.

Knight, L., & Edmonds, A. (2024). Understanding indirect meaning: A close replication. *Foreign Language Annals*, *57*(2), 450–471. https://doi.org/10.1111/flan.12719

Koike, D. A. (1989). Pragmatic competence and adult L2 acquisition: Speech acts in interlanguage. *Modern Language Journal*, *73*(3), 279–289. https://doi.org/10.1111/j.1540-4781.1989.tb06364.x

Koike, D. A. (1996). Transfer of pragmatic competence and suggestions in Spanish foreign language learning. In S. M. Gass, & J. Neu (Eds.), *Speech acts across cultures* (pp. 257–281). Mouton de Gruyter.

Language Teaching Review Panel. (2008). Replication studies in language learning and teaching: Questions and answers. *Language Teaching*, *41*(1), 1–14. https://doi.org/10.1017/S0261444807004727

Licea-Haquet, G. L., Velásquez-Upegui, E. P., Holtgraves, T., & Giordano, M. (2019). Speech act recognition in Spanish speakers. *Journal of Pragmatics*, *141*, 44–56. https://doi.org/10.1016/j.pragma.2018.12.013

Liu, F. (2025). Interpreting verbal irony in Mandarin Chinese: The role of prosody and proficiency among L2 learners of Chinese. *Journal of Pragmatics*, *249*, 120–138. https://doi.org/10.1016/j.pragma.2025.09.004

Marsden, E., Morgan-Short, K., Thompson, S., & Abugaber, D. (2018). Replication in second language research: Narrative and systematic reviews and recommendations for the field. *Language Learning*, *68*(2), 321–391. https://doi.org/10.1111/lang.12286

Martínez-Flor, A. (2008). The effect of an inductive-deductive teaching approach to develop learners' use of request modifiers in the EFL classroom. In E. Alcón-Soler (Ed.), *Learning how to request in an instructed language learning context* (pp. 191–225). Peter Lang.

McManus, K. (2022). Are replication studies infrequent because of negative attitudes? Insights from a survey of attitudes and practices in second language research. *Studies in Second Language Acquisition*, *44*(5), 1410–1423. https://doi.org/10.1017/S0272263121000838

McManus, K. (2023). How and why to conduct a replication study. In A. Mackey, & S. M. Gass (Eds.), *Current approaches in second language acquisition research* (pp. 334–351). Wiley-Blackwell.

McManus, K. (2024). The future of replication in applied linguistics: Toward a standard for replication studies. *Annual Review of Applied Linguistics*, *44*, 133–149. https://doi.org/10.1017/S0267190524000011

Meibauer, J. (2005). Lying and falsely implicating. *Journal of Pragmatics*, *37*(9), 1373–1399. https://doi.org/10.1016/j.pragma.2004.12.007

Nguyen, T. T. M. (2019). Data collection methods in L2 pragmatics research: An overview. In N. Taguchi (Ed.), *Handbook of SLA and pragmatics* (pp. 195–211). Routledge.

https://doi.org/10.1017/S0261444825101122 Published online by Cambridge University Press

Niezgoda, K., & Roever, C. (2001). Pragmatic and grammatical awareness: A function of the learning environment? In K. Rose, & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 63–79). Cambridge University Press.
Omar, A. (2006). Kiswahili requests: Performance of native speakers and learners. *Pragmatics and Language Learning*, *11*, 227–252.
Polio, C., & Gass, S. (1997). Replication and reporting: A commentary. *Studies in Second Language Acquisition*, *19*(4), 499–508. https://doi.org/10.1017/S027226319700404X
Porte, G. (2012). *Replication research in applied linguistics*. Cambridge University Press.
Porte, G., & McManus, K. (2019). *Doing replication research in applied linguistics*. Routledge.
Reins, L. M., & Wiegmann, A. (2021). Is lying bound to commitment? Empirically investigating deceptive presuppositions, implicatures, and actions. *Cognitive Science*, *45*(2), e12936. https://doi.org/10.1111/cogs.12936
Riddiford, N., & Joe, A. (2010). Tracking the development of sociopragmatic skills. *TESOL Quarterly*, *44*(1), 195–205. https://doi.org/10.5054/tq.2010.215252
Roever, C. (2005). *Testing ESL pragmatics*. Peter Lang.
Rose, K. (1992). Speech act research and written questionnaires: The effect of hearer response. *Journal of Pragmatics*, *17*(3), 49–62. https://doi.org/10.1016/0378-2166(92)90028-a
Sabaté i Dalmau, M., & Curell i Gotor, H. (2007). From 'Sorry very much' to 'I'm ever so sorry': Acquisitional patterns in L2 apologies by Catalan learners of English. *Intercultural Pragmatics*, *4*(2), 287–315. https://doi.org/10.1515/ip.2007.014
Safont-Jordà, M.-P., & Portolés-Falomir, L. (2013). Research method effects on third language learners' refusals. In O. Martí-Arnándiz, & P. Salazar-Campillo (Eds.), *Refusals in instructional contexts and beyond* (pp. 213–234). Rodopi.
Schauer, G. A. (2006). Pragmatic awareness in ESL and EFL contexts: Contrast and development. *Language Learning*, *56*(2), 269–318. https://doi.org/10.1111/j.0023-8333.2006.00348.x
Su, Y., & Chang, Y. (2019). Intra-lingual pragmatic variation in Mandarin Chinese apologies. *East Asian Pragmatics*, *4*(1), 59–86. https://doi.org/10.1558/eap.38215
Su, Y., & Ren, W. (2017). Developing L2 pragmatic competence in Mandarin Chinese: Sequential realization of requests. *Foreign Language Annals*, *50*(2), 433–457. https://doi.org/10.1111/flan.12263
Taguchi, N., Gomez-Laich, M. P., & Arrufat-Marques, M.-J. (2016a). Comprehension of indirect meaning in Spanish as a foreign language. *Foreign Language Annals*, *49*(4), 677–698. https://doi.org/10.1111/flan.12230
Taguchi, N., & Li, S. (2019). Replication research in contextual and individual influences in pragmatic competence: Taguchi, Xiao & Li (2016) and Bardovi-Harlig & Bastos (2011). *Language Teaching*, *52*(1), 128–140. https://doi.org/10.1017/s0261444817000222
Taguchi, N., Xiao, F., & Li, S. (2016b). Development of pragmatic knowledge in L2 Chinese: Effects of intercultural competence and social contact on speech act production in a study abroad context. *Modern Language Journal*, *100*(4), 775–796. https://doi.org/10.1111/modl.12349
Valdman, A. (1993). Replication study. *Studies in Second Language Acquisition*, *15*(4), 505. https://doi.org/10.1017/s0272263119000123
Warga, M., & Schölmberger, U. (2007). The acquisition of French apologetic behavior in a study abroad context. *Intercultural Pragmatics*, *4*(2), 221–251. https://doi.org/10.1515/IP.2007.012
Weissman, B., & Terkourafi, M. (2019). Are false implicatures lies? An empirical investigation. *Mind and Language*, *34*(2), 221–246. https://doi.org/10.1111/mila.12212
Wiegmann, A. (2023). Lying with deceptive implicatures? Solving a puzzle about conflicting results. *Analysis*, *83*(1), 107–118. https://doi.org/10.1093/analys/anac037
Winter, B., & Grawunder, S. (2012). The phonetic profile of Korean formal and informal speech registers. *Journal of Phonetics*, *40*(6), 808–815. https://doi.org/10.1016/j.wocn.2012.08.006

Cite this article: Bardovi-Harlig, K. (2026). Replication in L2 pragmatics. *Language Teaching*, 1–18. https://doi.org/10.1017/S0261444825101122

https://doi.org/10.1017/S0261444825101122 Published online by Cambridge University Press