



Towards Increased Classroom Assessment of Pragmatic Ability¹

Andrew D. Cohen²

Abstract

The article provides a rationale for assessing second language (L2) pragmatics in the classroom and then looks at tasks to assess comprehension of speech acts. Next we look at various ways to collect students' pragmatic production, such as through oral role-play, written discourse as if spoken, multiple-choice, or short-answer responses. Then six strategies for more effective assessment of pragmatics are presented, such as how to make the speech act situations realistic and how to rate for key aspects. The assumption made in the article is that classroom teachers may be avoiding the assessment of pragmatics, especially nonnative teachers who feel that they themselves are incapable of judging what constitutes correct behavior.

Keywords: *Pragmatic assessment, classroom-based teacher assessment, teacher-based assessment, L2 speech act performance assessment*

1. Introduction: Why Assess Pragmatics in the Classroom?

Tests of pragmatic ability have not tended to be part of classroom assessment. Even recently, the view taken is that “theories of communicative competence and communicative language teaching have not been fully developed and (that) rigorous empirical studies need to be carried out (first)” (Yamashita, 2008, p. 209). Given the importance of pragmatics in communicating in an L2, it would appear important to be assessing it in the classroom. Here are some of the reasons:

- Classroom assessment of pragmatics sends a message to the students that their ability to be pragmatically appropriate in the comprehension and production of language in different sociocultural situations is a positive thing.
- Having such items on a test motivates students to study about the performance of speech acts.
- It gives teachers an opportunity to see the relative control their students have in what may at times be a significant area for L2 performance.

¹ This article is a revised and updated version of Chapter 14 from Ishihara and Cohen, 2010.

² Professor Emeritus, University of Minnesota, USA. E-mail: adcohen@umn.edu



- It gives teachers an opportunity to see whether learners have learned the pragmatics that were included in the instruction.

While there should be a good fit between what is taught and what is assessed, in practice it does not always the case since instructional content does not always lend itself to easy assessment. To make this more concrete, take the speech act of apologizing. How do we determine a standard for correctness? Is it what natives do or are supposed to do? Isn't it possible that natives won't necessarily get it right, as in the case, say, of a husband's half-hearted apology to his wife that fails to appease her? For example, let us say that the husband says with a slightly facetious tone of voice, "Sorry about that, dear." The wife may well respond with "Well, I'm still upset!" But if the husband were to say with a sincere tone, "I'm really sorry about that, darling," he would more likely appease her.

So, pragmatic ability is challenging to measure. First, teachers need to make sure that the context is clear since situations differ in terms of what the appropriate answers will be. And how is the score for performance determined? Depending on the situation, even highly competent speakers of the given language may differ in their responses such that it is not clear what an acceptable or unacceptable answer would be (McNamara &Roever, 2006). Consequently, keeping the test practical in administration and scoring may not be easy. For instance, role-plays take time to conduct and may require multiple ratings. In addition, it may be a challenge to compare one with another. Also, constructing a discourse completion test (DCT) can be a challenge, especially given the need to make sometimes minute distinctions for testing purposes (McNamara &Roever, 2006, pp. 54-75). Thus, it can take creativity to design instruments that test pragmatics reliably and without expending an excess of time and effort (Roever, 2004). Brown (2008), for instance, took a close look at statistical analysis for five of the common pragmatics measures. His findings led him to recommend using a generalizability theory approach to statistical analysis of pragmatics data, which might cause more problems for those with a limited statistical background than would using a classical theory approach.

To add to the complexity, a study demonstrated systematic interrater differences in L2 pragmatics ratings for three trained raters across three test types (an open-ended written discourse completion task, a language lab DCT, and a role-play), three speech acts (refusals, apologies, and requests), two proficiency levels of respondents (intermediate and advanced learners of Korean as a foreign language), and specific test items (Youn, 2007). The results of the study indicated that all three raters showed different degrees of severity in their ratings, depending on the test type and speech act. In addition, the raters showed unique bias patterns in their ratings of given respondents.

Despite the challenges posed by attempting to assess complex speech actbehavior, there are enough largely routine patterns among native or near-native speakers of the language to warrant both teaching these to learners and assessing their approximation of them. There are, in fact, benefits to including pragmatics among those language features that are assessed, especially if comparing key differences between two language communities. The emphasis in teaching would be on situations where divergence from the native-like norm in performance may lead to *pragmatic failure* (that is, where the result or uptake from the speech act interaction is not what the speaker desired, usually because the speech act somehow violated the norms of the speech community).



If we as teachers agree that pragmatics has an important role to play, then we can support that position by being sure to include its assessment in both short quizzes and longer tests. It goes without saying that learners pay extra attention to what they are going to be tested on. So let us say, for example, that an English-as-a-foreign-language (EFL) teacher provides instruction on how to complain effectively in English (e.g., a customer to a sales clerk about the slow service in the store). It would then be useful to test for the students' ability to perceive and produce such complaints.

A qualitative case study was conducted in a freshman English course at a Japanese university in order to demonstrate how pragmatics could be integrated into a classroom context and to demonstrate the effectiveness of teacher-based assessment of pragmatic competence grounded in Vygotsky's sociocultural theory (Ishihara, 2009). The teacher researcher implemented pragmatics-focused instruction throughout a semester based on empirically-established information on speech acts in English. The instructor used collaboratively-developed, authentic assessment tools, such as reflective writing, rubrics, role-plays, and self/peer-assessment, and facilitated interaction and assessment in the learning process.

In the current article, we will explore some principles for constructing tasks that assess learners' pragmatic comprehension and production, and then consider some of the more frequently used approaches for doing so. The article then provides six strategies for more effective assessment of pragmatics, and includes two suggested activities in the appendix, one for assessing comprehension and the other production.³

2. Approaches to Assessing L2 Speech Act Performance

In this section we will look both at how to elicit speech act performance and also how to evaluate the students' responses once they have been elicited. A prime concern is making sure that your students understand the situations in which pragmatic performance is being assessed. It may be appropriate, for example, to describe the situation in the students' first-language (L1), in order to make sure that they fully understand the contextual factors involved. Such factors might include among other things: (1) the degree of imposition in a request or the severity of the infraction in an apology situation, (2) the level of acquaintance between the speaker and hearer, and (3) their relative social status.

A caveat with regard to the examples of pragmatics in this article is that they generally adhere to the more traditional approach of looking at speech acts in isolation. While this approach may be beneficial for learners attempting to understand how speech acts actually function, Kasper (2006) and more recently Roever (2011) have argued that this approach under-represents the construct of pragmatic ability. Roever contends that tests need to include real-time measures of learners' interactional abilities to allow defensible extrapolation to a target domain of social language use. Roever (2011) notes in particular, that the assessment of learners' ability to produce extended monologic and dialogic discourse is a missing component in existing assessments. Roever prefers, for this reason, the use of open role plays (an approach which is addressed below).

2.1. Measuring Comprehension of Pragmatics

³ For a more research-based approach to assessing pragmatics, see Roever, 2004; Cohen, 2004.



Assessment of pragmatics looks both at comprehension and production. Measuring *comprehension* of the pragmatics in language behavior often takes an indirect, metapragmatic approach. Learners are asked to assess how well they **think** someone else has performed pragmatically. So, for instance, assessment could entail looking at learners' reactions to videotaped role-play, screen plays (from TV series), or written descriptions of speech act situations. The learners indicate their reactions by completing rating scales or multiple-choice items. So, taking the issue of marital relations, let us suppose that your students viewed a video clip which had the following exchange between a husband and wife:

Example 1

Wife: I don't like it, dear, when you criticize our children in front of other people. It made me uncomfortable last evening when you criticized them at the dinner party. I know you were trying to be funny, but people can take it the wrong way, and...

Husband: Really? I don't agree with you. In fact, I think you're overreacting – it's not such a big deal. But if you insist, I'm willing to watch what I say....

A sample item could include the following:

Indicate with an X how you would rate the level of the husband's apology:

1. High ____
2. Moderate ____
3. Low X
4. Non-existent ____

In addition, learners of the particular target language (in this case, EFL) could be asked to add their rationale for giving a certain rating. So in the above example, a given learner might add, "It isn't apologetic enough. The tone is sarcastic and like criticizing his wife. It isn't clear that he will actually change his behavior."

Learners may have differing reactions, many of which could prompt class discussion. If the focus is on the likely effect of the speaker's delivery of a speech act on the addressee (i.e., the *uptake*), then we could start by assessing how the speech act is perceived by the learners. The following is a sample vignette where the students are to respond both in a multiple-choice format and also provide an open-ended comment as well:

Example 2

George is doing his holiday shopping in Manhattan and has only about 15 minutes before the department store closes. He needs to get across the entire store to the opposite corner to check out the gift specials at the men's accessories counter, but in front of him is a rather obese lady with bags in hand. She is in the midst of a heated conversation on her cell phone and is blocking



the aisle. George tries to get around her, but in the process inadvertently knocks over some of her bags, tangles up her cell phone arm, and causes the lady to drop her phone as well.

Lady: My goodness! What are you doing, young man?

George: Very sorry, lady, but you were in my way!

How likely is the stout lady to consider George's response an apology?

- a. Very likely.
- b. Somewhat likely.
- c. Not very likely.
- d. Not at all likely.

What is your rationale for your choice? _____

The teacher may have taught the students how the intensifier "really" in an apology would signal regret and "very" would be more a marker of etiquette. Hence, the students should have answered "c" or "d" as the correct multiple-choice response. In this case, the strategy of "acknowledging responsibility" is the key concern here. A preferred rationale statement would be as follows:

George isn't really taking responsibility for knocking into the lady. He's putting the blame back on her.

Tasks like this one would be intended to tap the learners' understanding of speech act performance by others, and in this case, the added ingredient of annoyance or sarcasm. Sometimes the identification of the speech act is not so challenging, but its function in the interaction can depend on the tone or attitude of the person who delivers it.

Assessment of the ability to comprehend pragmatics could also entail measuring it directly. For example, teachers could construct some items for their learners which have pragmatic input that is perhaps subtle in nature (e.g., veiled criticism about a talk the person just gave or a complaint about the person's working style). The intent would be to see if the learners understand the illocutionary force of the message and are able to deal with it in an appropriate way pragmatically.

2.2. Measuring Pragmatic Production

2.2.1. Oral Role-Play

With respect to *oral production*, we can observe students engaged in spontaneous speech, although it is not easy to collect such data. Another problem with spontaneous speech is that it may not be comparable across students. Some students say more than others and are thus likely



to provide more language in their oral speech act performance than are other students. Also, learners may be having a bad day or may have misunderstood the given testing task.

When eliciting samples of speech, it is helpful to give the students some warm-up time, rather than to assume that they are ready to perform pragmatically on request. Their minds may need to get going in the target language first. So it helps to give students a chance to rehearse what it is they will need to say or write. If it is to be oral, they could even be given time to rehearse it with a partner.

Given the problems with assessing spontaneous speech, it is preferable to set up role-play situations, such as the following:

Example 3

You completely forget a crucial meeting at the office with the boss at your new job. An hour later you show up at his office to apologize. The problem is that this is the second time you've forgotten such a meeting in the short time you have been working at this job. Your boss is clearly annoyed when he asks, "What happened to you this time?" Do your best to smooth over the situation.

The sociocultural factors of interest in measuring pragmatic ability, such as in a role play like this, often include the following:

- the relative status of the speaker and hearer,
- the level of acquaintance of the speaker and hearer,
- the degree of severity, imposition, or other impact caused by the speech act situation.

In this case, the speaker is clearly of lower status in being an employee. Since it is a relatively new job, the speaker is probably not that acquainted with the boss. The meeting was crucial so the severity level is high. The responses in such role-play situations could be rated for:

- the ability to use an appropriate speech act,
- the typicality of the expressions used,
- the appropriateness of the amount of speech and information given,
- the appropriateness of the level of formality, the directness, and the level of politeness.

So let us say the response from a nonnative speaker of English looks like this:

Very sorry, Mr. Iverson. You see...uh...I have sleeping problems and...uh... then I missed the bus. But I can make it up to you.

In this instance, the speaker has apologized using the strategies of expressing an apology, giving an explanation, and offering repair. So three of the key apology-specific strategies were used. Nonetheless, it is likely that pragmatically competent speakers of American English would



not state the explanation and the offer of repair that way. They would probably say something more like the following:

Oh, I'm really sorry about that, Mr. Iverson. I've been suffering from chronic sleep disorder and as a result I have trouble getting going in the morning. I can get you a doctor's note about it. And to make matters worse, I got to the bus stop this morning just as the bus was pulling away. I'm really sorry about that. What can I do to make it up to you? I'll work overtime, whatever.

Let us assume that the excuse here is legitimate, both with regard to the sleeping disorder and the missed bus. A sympathetic boss in a U.S. company would probably accept it, especially with a doctor's note. Notice that native-like performance would probably entail going into a little more detailed explanation of the actual health problem (e.g., giving it a label) and of what happened with the bus (rated under point "c" above). In addition, such performance would also probably include suggested ways to make it up to the boss, such as by working overtime.

Here is another possible response from the nonnative office worker:

So sorry I missed the meeting. I had problem at home and then I forgot the meeting and when I remembered it was too late.

In this instance, the worker expresses an apology, gives an explanation, and acknowledges responsibility. As in the previous response by a learner, the speaker used three strategies for apologizing. Is the response apologetic enough? Most English speakers might express the apology that way. Would the boss accept it? In this instance, the average boss would probably point out that this has happened before and that the worker needs a better system for remembering meetings.

The number of strategies called for in performing a given speech act will vary according to the situation and speech community. For example, the strategy of "expressing an apology" may be sufficient for setting things right in an e-mail exchange with a colleague in a U.S. academic institution: "I'm really sorry for that inappropriate e-mail I sent you." Likewise, the strategy of offering an explanation or an excuse, "The bus was late," when an employee arrives late for a meeting in the Middle East, may be accepted by the boss as sufficiently apologetic, since busses can be late there and consequently employees are only partially responsible for getting to meetings on time.

The following are two further examples of oral role-play prompts. Here is the first:

Example 4

Your next-door neighbor keeps his dog out on his porch well into the evening and the barking is driving you crazy. Role-play the part of the irate resident who knocks on the neighbor's door and requests that the dog be kept inside at night. Your partner will play the role of the partially deaf, elderly neighbor, Grace Smith, who wants to keep this dog happy since it was the pride and joy of her deceased husband.

Elderly neighbor, Grace: Well, hello, _____. What can I do for you?



You:

Here is the second example of an oral role-play prompt:

Example 5

You promised your classmate friend that you'd get tickets in advance for a special showing of a movie but forgot, and now the show is sold out.

Classmate: What a bummer! I really wanted to see that movie this evening. I was supposed to report on it in film class tomorrow.

You: _____

The first situation calls for being tactful and understanding, while making the need for quiet very clear. Presumably the EFL course has introduced tactful requests, and the rating here would be for the ability to make a polite request. The second situation calls for a profuse apology, and a convincing one. The offer of repair here may be crucial (e.g., offering to drive the classmate to a more distant theater where tickets are still available). Open-ended speech production in situations like these requires that the learners perform a sometimes challenging search of their memory and then select the appropriate forms from a wide array of possible solutions (Kasper, 1999). Also, open role-plays approximate authentic interactions in that there is the full operation of turn-taking, sequencing of moves, and negotiation of meaning (Kasper & Dahl, 1991).

In a cross-sectional design aimed at viewing the speech act of requests embedded in discourse, Al-Gahtani and Roeber (2012) collected open role-play data from learners at four proficiency levels, and focused on the sequential organization of the interactions and the impact of participants' proficiency level. Findings indicated that lower-level learners were less likely to project the upcoming request and lay the groundwork for it through ascertaining interlocutor availability and providing accounts. In addition, they were less likely to supply the first utterance in an adjacency pair and made the request early, relying on the interlocutor to elicit further information.

2.2.2. Written Discourse as if Spoken

An alternative to assessing oral language production would be to have the students produce written responses, intended to reflect what they would say in the situation. Although it is a written test, it could still elicit learner language efficiently. Actually, learners may be able to provide more thoughtful or socially desirable responses in such written tests, possibly even more indicative of their knowledge of what a speaker might say than when having to respond orally (Beebe & Cummings, 1996; Kasper & Dahl, 1991).

The DCT, mentioned above, is a popular approach to the assessment of speech acts in written discourse. The format often calls for two or more turns by both the speaker and the hearer that the student needs to take into consideration in responding:



Example 7

You arranged to meet a friend in order to study together for an exam. You arrive half an hour late for the meeting.

Friend (annoyed): I've been waiting at least half an hour for you!

You-1:

Friend: Well, I was standing here waiting. I could have been doing something else.

You-2:

Friend: Well, it's pretty annoying. Try to come on time next time.

Here are responses to You-1 and You-2, provided by a native speaker of Hebrew:

You-1: So what! It's only an – a meeting for – to study.

You-2: Yeah, I'm sorry. But don't make such a big deal of it.

If the responses were being rated for the ability to be properly informal with a friend, then this respondent would come out high on that scale. The responses are informal. On the other hand, this reply is likely to be seen by the friend as not very apologetic and might get rated down on that basis.

Here is another apology DCT, in which the severity of the infraction is major, the learner respondent as hearer in this situation is highly acquainted with the speaker (being the professor's intern), but the learner's relative status as student is considered lower:

Example 8

You are a graduate assistant for a professor who requested that you pick up a library book to help him finish the review of literature for a research proposal which is actually due tomorrow. You arrive at the meeting without the book. The problem is that you were supposed to get the book to him at your meeting last week and it slipped your mind then.

Professor: Do you have that book we need in order to finish up the review of literature?

You-1:

Professor: Yeah, but you actually said you were going to get it for our meeting last week, and you didn't bring it then either.

You-2:



Professor: Still, I think you might need a better system of tracking your tasks as my RA. We have to get this finished today so we can submit it tomorrow.

You-3:

Professor: OK. I guess that'll help but I'll be teaching when you get back so you'll have to work through that section on your own and leave your suggestions on my desk.

You-4: _____

The query by the professor would presumably be answered by some explanation or excuse and then by an expression of apology with the appropriate intensity attached (*Y-1: Whoops! I forgot it. You can't imagine how many things I've had on my mind lately... I'm really sorry about that.*). The professor's reply might be met by a further expression of apology (*Y-2: Oh, well, I'm terribly sorry about that...*). The next turn calls for an offer of repair (*Y-3: I'll go and get the book right now.*). But as often happens in life, the professor is no longer available at that point to work on the task. The graduate assistant would then need to apologize again and perhaps offer a promise of non-recurrence (*Y-4: I'll make sure nothing like this happens again.*).

As indicated above, if the person doing the scoring of the responses knows which apology strategies are likely to be used in each instance, then scoring the responses becomes easier. The use of multiple turns in a DCT task is intended to reflect an actual interaction more than is possible with a single prompt, a space for a reply, and no subsequent turns by the addressee. A way to score the above interaction could be to give a "3" if the response seems fine (which is probably the case for the responses offered in the preceding paragraph), "2" if it is fair, and "1" if it is weak for each of the four slots. So, 12 points would be the highest possible score on the item. The score could take into account the following five elements in a holistic fashion:

- the typicality of the expressions used,
- the appropriateness of the amount of speech and information given,
- the appropriateness of the level of formality,
- the directness,
- the level of politeness.

The development of tasks such as these allows us to manipulate the social factors (such as status, as in student-professor interactions) and situational factors (such as the level of imposition or of severity in an apology situation). The use of multiple turns represents an effort to make the DCTs more reflective of the conversational turn-taking that takes place in actual speech, the lack of which has been a criticism of written DCTs (Bardovi-Harlig & Hartford, 1993). This multiple-rejoinder format makes the students structure their speech act behavior to conform to a given situation. It helps to indicate in the instructions that respondents are to respond according to what people would typically say. This then would give the instructor a sense of how aware they are as to the generally preferred or commonly used target-language behavior.

2.2.3. Multiple-Choice and Short-Answer Completion Items



Teachers can also assess oral production indirectly, such as by using multiple-choice items since they are easier to score than are open-ended items. Such items may seem like measures of how well they comprehend the pragmatics of the situation, but since they are intended to tap the ability to produce the correct responses, they could be viewed as doing it as a measure of “projected ability.” Consider the following:

Example 9

Indicate which of the following is most likely Bill's response to Andrew in leave taking.

Andrew: Hey, Bill. It's been nice talking with you. Let's get together some time.

- Bill:*
- a) Good idea – when would you like to do it?*
 - b) You always say that but don't mean it.*
 - c) Sounds good. Take care.*
 - d) I won't hold my breath.*

It is likely that for native speakers of American English, "c" would be the expected choice in that this is most likely what people might say. The other three response choices are possible, especially "a," but selecting them would mean departing from the expected speaking routines for the average speech community. The vague statement "Let's get together some time" does not usually constitute an actual invitation.

Recent research used verbal report techniques to describe the 24 different strategies that German respondents used to answer multiple-choice items on a comprehension measure of American English pragmatics (Timpe, 2012). The study demonstrated that respondents with more exposure to the target-language context had an easier time contextualizing the items than those with less exposure, who had to rely more on the text of the items. With regard to the multiple-choice format itself, the verbal report data confirmed that while numerous items were assessing the intended sociopragmatic aspects of the situations, some of the items were not.

Another approach to assessing production of pragmatically appropriate forms would be to use a sentence completion format:

Example 10

Herman is requesting a raise from his boss. Complete his request so that it sounds tactful:

I was _____ if you _____ consider increasing my pay.

It is possible through this format to see if the students have control over certain grammatical and lexical forms that are routinely used in order, for example, to mitigate or soften their requests.

The following would be an example of a "guided" response situation if the learners were provided, say, with the base form of the lexical item to use:



Example 11

I was _____ if you _____ consider increasing my pay a bit.
(to wonder) (will)

Items of this sort might precede the ones where no base form is given as a clue to the intended response for the given blank.

3.Suggested Strategies for Assessing Pragmatics

Let us conclude this article by considering six strategies for assessing pragmatics. Then we provide two activities to which these strategies can be applied.

3.1. Keep the speech act situations realistic (for the learner group) and engaging

So if the students are learning EFL, you could include a vignette about babysitting since this is likely to be a culturally acceptable activity in, say, the U.S. since family members are not usually called upon to do the babysitting. In terms of finding vignettes that are likely to occur and engaging, the logical approach would be to check with locals of the given speech community. But another source can be the learners themselves since they are the ones who may be acutely aware of just those situations for which they would like guidance in pragmatics (McKlean, 2005). Here is one such situation supplied by aEFL student in Japan:

You are at a restaurant and someone at your table says something funny. You laugh and spray a little food. You are embarrassed and think you should apologize. What do you say? (McKlean, 2005)

In other words, students may be excellent at furnishing situations, regardless of whether they have any idea as to how to deal with them in a pragmatically appropriate way. So, for example, the teacher makes the following request:

Think of a social situation where you had problems communicating (e.g., making a request or refusal, apologizing for something you did, complaining about something). Describe the situation briefly, and end it with "What do you say?" (McKlean, 2005)

3.2. Check for key aspects of performance

a) the cultural appropriateness of the strategies in the given situation. So, the students need to rate the following two questions addressed to co-workers at the office in terms of their cultural appropriateness:

"I see you got a new car. How much did you pay for it?"

"How much do you make a month?"



The performance could be scored on, say, a broad 4-point scale: 'totally appropriate culturally,' 'somewhat culturally appropriate,' 'somewhat culturally inappropriate,' 'culturally inappropriate.'

- b) the appropriateness of the language forms used with regard to the level of formality (e.g., "too informal," "just right," or "too formal"), the degree of politeness (e.g., whether appropriate for that situation in that language and culture; 4 – appropriate, 3 – somewhat appropriate, 2 – somewhat inappropriate, 1 – inappropriate), and the amount of language used (e.g., "too much," "just right," or "too little").

The best rule of thumb is for teachers to be generous in their ratings since there is variation among native speakers.

3.3. Have a discussion with the students after they have performed speech acts

Ask the students how well they understood the contextual factors, and have them identify the factors that most contributed to their responses.

3.4. Have the students compare their performance with that of a native

It could be helpful to have the students respond to these questions:

- a) What do they think a native-like response would be?
b) How do they think their own L2 pragmatic performance might depart from the native norms if they are unwilling to do it the way natives would.

This allows for a discussion of the how the students relate to the pragmatics of the L2 speech community (see Ishihara, 2006). So in the situation of forgetting a meeting with their boss for the second time, students of Hebrew could be alerted by their teacher to the likelihood that native Hebrew speakers might well express an apology **without** offering a repair. Then if the students were expressing their own subjectivity as, say, an American speaking Hebrew, they might indicate that they would offer repair (e.g., "I can be there in 5 minutes and deal with the matter."). While the score in this situation would be based on the awareness of the community norms, the teacher might not penalize the students for departing from those norms **if** they are able to state it as a conscious preference.

So students would be rated **both** for their awareness of the norms (e.g., 3 – high awareness, 2 – moderate awareness, 1 – low awareness) **and** for their explanation for why they depart from or resist that norm (e.g., 3 – clear statement of reasons for resistance, 2 – some statement of reasons for resistance, 1 – no statement of reasons for resistance). If the students write that native speakers would not offer repair (because it would be viewed as yet another infraction given their status as employee), they would get a "3" for being highly aware of the cross-cultural difference. If they go on to note that for them it would be imperative to indicate to the boss what they were willing to do to make it up, they would get a "3" for noticing the gap between what was expected and their own preferred behavior.



3.5. Have the students provide a rationale for why they responded as they did in the given social situation.

If the group is not too large, their reactions could be recorded digitally at the time they are responding (if in writing) or afterwards. Otherwise, they could provide a brief written rationale, which would actually facilitate evaluation of their pragmatic understanding. The request for providing a rationale could look like this:

"Please give a brief explanation for why you responded in this situation the way you did. What were the factors that influenced your response?"

3.6. Be strategic about when to assess what

At the beginning of the course and with less proficient learners, less complex speech acts could be assessed, like complimenting, thanking, leave taking, and basic service encounters (e.g., simple requests). While there is a need to teach (and assess) some aspects of requesting, making invitations, refusing, apologizing, and complaining early on, the more complex aspects of such speech acts could be reserved for assessment later in the course or in courses for more advanced levels of instruction.

The appendix of this article offers two activities intended to provide an opportunity for constructing at least one task involving comprehension of a speech act and one task involving production of a speech act. The activities include both suggested situations and possible formats for the tasks, with the instructions either in the target language exclusively or bilingually.

4. Conclusions

It would appear that even if pragmatics is taught in the EFL classroom, it is not necessarily assessed. When learners get the pragmatics wrong, this can lead to pragmatic failure, which may mar the quality of interactions. The article has given some examples of how to conduct assessment of pragmatics and some suggestions as to what to look for in responses. While collecting pragmatics data from students is a challenge, even more of a challenge is determining how to score the responses, taking into account the degree of imposition or severity of an infraction, the level of acquaintance, and the relative social status that is implied in the given situation. Examples were given of how to open up the assessment to provide respondents an opportunity to express their take on the situation and to explain why they responded as they did. The process of assessing pragmatic ability offers numerous opportunities for research since we are still learning how best to conduct these assessments in the classroom. This article has given numerous suggestions as to possible areas for research:

- Conducting qualitative studies of what constitutes effective pragmatics instruction in the classroom.



- Identifying the most effective means for assessing speech acts and other pragmatic behavior in a full, interactive discourse context, whether through open role-play or other means.
- Continuing to collect data on the strategies reportedly used in responding to multiple-choice pragmatics items so as to better understand what is being assessed.
- Continuing research on inter-rater agreement and disagreement in assessing pragmatic ability.
- Finding statistical approaches to analyzing pragmatics data that are both easy to use and to interpret.

Hopefully future research will inform us as to how best to streamline such assessments in the years to come.



References

- Al-Gahtani, S. & Roever, C. (2012). Proficiency and sequential organization of L2 requests. *Applied Linguistics*, 33(1), 42–65.
- Bardovi-Harlig, K. & Hartford, B. S. (1993). Refining the DCT: Comparing open questionnaires and dialogue completion tasks. *Pragmatics and Language Learning*, 4, 143–165.
- Beebe, L. M. & Cummings, M. C. (1996). Natural speech act data versus written questionnaire data: How data collection method affects speech act performance. In S. M. Gass & J. Neu (Eds.), *Speech acts across cultures* (pp. 65–86). Berlin: Mouton de Gruyter.
- Brown, J. D. (2008). Raters, functions, item types and the dependability of L2 pragmatics tests. In E. Alcón Soler & A. Martínez-Flor (Eds.), *Investigating pragmatics in foreign language learning, teaching and testing* (pp. 224–248). Bristol, England: Multilingual Matters.
- Cohen, A. D. (2004). Assessing speech acts in a second language. In B. Boxer & A. D. Cohen (Eds.), *Studying speaking to inform second language learning* (pp. 302–327). Clevedon, England: Multilingual Matters.
- Cohen, A. D., Olshtain, E., & Rosenstein, D. S. (1986). Advanced EFL apologies: What remains to be learned? *International Journal of the Sociology of Language*, 62(6), 51–74.
- Ishihara, N. (2006). "Centering second language (SL) speakers' experience: A study of SL speakers' resistance to pragmatic norms of the SL community," in *Subjectivity, second/foreign language pragmatic use, and instruction: Evidence of accommodation and resistance*. Unpublished doctoral dissertation (Ch. 6, pp. 162–182). Minneapolis, MN: University of Minnesota.
- Ishihara, N. (2009). Teacher-based assessment for foreign language pragmatics. *TESOL Quarterly*, 43(3), 445–470.
- Ishihara, N. & Cohen, A. D. (2010). *Teaching and learning pragmatics: Where language and culture meet*. Harlow, England: Longman Applied Linguistics/Pearson Education.
- Kasper, G. (1999). Data collection in pragmatics research. *University of Hawai'i Working Papers in ESL*, 18(1), 71–107.
- Kasper, G. (2006). Speech acts in interaction: Towards discursive pragmatics. In K. Bardovi-Harlig, C. Félix-Brasdefer, & A. Omar (Eds.), *Pragmatics and language learning* (Vol. 11), pp. 281–314. HI: University of Hawai'i at Manoa.
- Kasper, G. & Dahl, M. (1991). Research methods in interlanguage pragmatics. *Studies in Second Language Acquisition*, 13(2), 215–247.
- McNamara, T. & Roever, C. (2006). *Language testing: The social dimension*. Malden, MA: Blackwell.
- McLean, T. (2005). "Why no tip?": Student-generated DCTs in the ESL classroom. In D. Tatsuki (Ed.), *Pragmatics in language learning, theory, and practice* (pp. 150–156). Tokyo: Pragmatics Special Interest Group of the Japan Association for Language Teaching.
- Roever, C. (2004). Difficulty and practicality in tests of interlanguage pragmatics. In B. Boxer & A. D. Cohen (Eds.), *Studying speaking to inform second language learning* (pp. 283–301). Clevedon, England: Multilingual Matters.
- Roever, C. (2011). Testing of second language pragmatics: Past and future. *Language Testing*, 28(4) 463–481



- Timpe, V. (2012). *Strategic Decoding of Sociopragmatic Assessment Tasks — An Exploratory Think-aloud Validation Study* *Second Language Studies*, 30(2), 109–246.
- Yamashita, S. (2008). Investigating interlanguage pragmatic ability: What are we testing? In E. AlcónSoler & A. Martínez-Flor (Eds.), *Investigating pragmatics in foreign language learning, teaching and testing* (pp. 201–223). Bristol, England: Multilingual Matters.
- Youn, S. J. (2007). Rater bias in assessing the pragmatics of KFL learners using facets analysis. *Second Language Studies*, 26(1), 85–163.



Appendix

Activity 1: Writing a Task to Assess Comprehension of Speech Acts

Format: Individual, pair, whole group

Objectives:

- 1) To give you an opportunity to construct a task that checks for learners' ability to appraise a speech act situation..
- 2) To select from among various formats in the construction of these speech act comprehension tasks.

Time: 1 hour

Materials:

- Four situations and two formats for the task construction: *Situations and Formats for Writing a Speech Act Comprehension Task* (see below).

Directions:

- 1) Select a partner for this activity.
 - 2) With your partner, start by reviewing the two speech act comprehension tasks presented above (the one between the husband and wife, and the bumping accident in the department store) with regard to how comprehension of speech act performance is assessed.
 - 3) Jointly select your item formats for assessing comprehension of pragmatic ability – such as:
 - a) multiple-choice,
 - b) open-ended questions (to probe students' rationale for their choice of multiple-choice response), and
 - c) rank-ordering of responses (such as for politeness).
 - 4) Select a situation (four are provided below, but you could use your own if you prefer).
 - 5) Determine the criteria that you wish to use in your assessment (e.g., degree of regret in keeping with the nature of the infraction, sincerity of apology).
 - 6) After you have blocked out your task, take turns responding to it, whether in English or in an L2 that you both know. This will serve as a form of a “pilot” for your task. If possible, audio-record your responses so that you can re-listen to them to facilitate analysis.
 - 7) Each pair is to report back to the whole group as to what the task is intended to measure, how learners' responses are to be scored, the relative ease in constructing such a task, and potential challenges associated with it.

Discussion/Wrap Up: Look at ways that this exercise has helped sharpen your understanding of how pragmatic tasks can be constructed.



Situations and Formats for Writing a Speech Act Comprehension Task

1. Situations

- a. Forgetting to get the medicine you promised to pick up at the pharmacy for the sick child of a single-parent neighbor.
- b. Apologizing to your spouse for accidentally revealing personal information about her at a dinner party – information that she didn't want revealed.
- c. Apologizing to a college student for causing him to have to swerve on his bike due to your jaywalking.
- d. Apologizing to your mother for forgetting to send her a birthday card.

2. Formats

- a. Multiple-choice: You could make up a multiple-choice response to a situation where learners have to choose the most appropriate response for a given situation.
- b. Rank order: You could have your students rank the order of responses (e.g., in terms of level of directness). You give a series of apologies which are on a continuum from more casual and detached to more regretful and engaged (e.g., "Oh, sorry about that..." to "Oh, my goodness, I am really very sorry about that...").
- c. Open-ended questions: In this case, since it is for comprehension and not production, the use of open-ended questions would be, say, to probe students' rationale for their choice of multiple-choice response. Note that this is not a format for assessing comprehension directly, but rather as a means for probing cognition – to better understand the response process.

Activity 2: Writing a Task to Assess Production of Speech Acts

Format: Individual, pair, whole group

Objectives:

- 1) To give you an opportunity to construct a task that has learners perform orally or in writing their ability to engage in a speech act interaction.
- 2) To establish criteria for judging the effectiveness of the given speech act performance.



Suggested time: 1 hour

Materials:

- Three situations and three formats for constructing a task: *Situations and Formats for Writing a Speech Act Production Task* (see below).

Directions:

- 1) Start by reviewing the speech act production tasks in this article (four apologies: forgetting to meet with the boss, forgetting to buy movie tickets, forgetting a study date with a friend, and forgetting to get a library book for a session with a professor; a complaint: about the neighbor's dog).
- 2) Select a format for assessing production of pragmatic ability from Section #2 of the worksheet below.
- 3) Select a situation, either from the list of suggestions or from your own experience.
- 4) Determine the criteria that you wish to use in your assessment (e.g., appropriateness of directness or politeness given the seriousness or severity of the situation with regard to a complaint, request, or apology).
- 5) After you have constructed a draft version of your assessment task, find a partner and have that person respond to it, whether in English or in an L2 that you know. This will serve as a form of a “pilot” for your task. Pay attention to similarities and differences in the written answers that you produced, in an effort to be mindful of variation in responses (especially if you are both competent speakers of the language).
- 6) Each pair is to report back to the whole group as to what the task is intended to measure.

Discussion/Wrap Up: Look at ways that this exercise has helped sharpen your understanding of how pragmatic tasks can be constructed. The discussion could include participants’ comments about the relative ease involved in constructing such a task, potential challenges associated with it, and ways to address these challenges.

Situations and Formats for Writing a Speech Act Production Task

1. **Situations** – either using the example provided or creating your own

- a. Requesting a second opinion (e.g., from another doctor at the hospital when the diagnosis for your loved one’s situation is unacceptable to you).



- b. Complaining (e.g., to the manager of a pricey restaurant manager about slow service and mediocre food).
- c. Complimenting (e.g., a very modest friend in front of others on the good advice that friend gave you).
- d. Apologizing (e.g., to a female colleague for your unpleasant behavior towards her at a meeting the previous day. You weren't feeling well and took it out on her.).

2. Formats

- a. *Guided response: Multiple-rejoinder response.* You could use this format of providing three or four responses that the hearer provides in the interaction and leaving the answers open for the learner to complete (as in the apology to the professor for forgetting the library book).
- b. *Open-ended response: Oral role-play response.* You design a situation with a description such as in forgetting the meeting with the boss for the second time (see above). Possibly you provide an initial utterance from the speaker (in this case, the boss) in the interaction, as in "What happened to you this time?"
- c. *Open-ended response: Written.* The format is like in "b" above, but in this case the learner writes a response. It would most likely be in the form of a single answer.