
Web-based evaluation of EFL learners' pragmatic competence utilizing video stimuli

Nicholas O. Jungheim

Key words: language testing, pragmatics, speech acts, second language acquisition, EFL

[Abstract] Pragmatic competence has long been an element in a wide range of frameworks for communicative competence. Considering the options for testing pragmatic competence, however there are still problems related to authenticity, validity, reliability, and ease of administration. Thus, this report describes the development of interactive online test as a way to overcome some of the limitations of previous tests of pragmatic competence. The e-Prag test is an Internet-based means of providing visual stimuli using video clips to collect video data from Japanese EFL learners performing a series of speech acts in authentic situations. The overall theme of the test is speech acts for university life. It is designed to test Japanese learners' ability to perform a limited number of speech acts that they could use even in the context of a Japanese university campus. In the first stage, a series of five videos were taken to elicit an apology, a response to a complement, a request for help, a suggestion, and a refusal of help. The scenarios varied according to degrees of power, social distance, and imposition. After arranging the videos in a PowerPoint presentation, they were administered to 11 Japanese university students and three native English speakers whose responses were recorded with a video camera. In the second stage, the videos were rearranged in html format for administration over the Internet, with some revisions based on input from the pilot study participants. Furthermore, the online version includes a sample item, sample answer video clips for a built-in washback effect, and a form at the end for the test takers to input their personal information and any questions or comments that they might have. In the yet to be carried out third stage of development, the e-Prag test will be administered to a group of Japanese learners of English in order to gather more sample responses to fine-tune the rating scales and create a training video to help raters calibrate their ratings. Finally, a large-scale administration will be carried out to validate the test by comparing the results with those of other measures of pragmatic competence. This project succeeded in developing a workable, easy-to-administer online test in the limited amount of time allotted for the study and with the limited resources available, but further programming is required before it is possible to upload test takers' videos to a server.

[Introduction] There is no question that pragmatic competence plays an important part in language learners' communicative competence, and it has been included in various models of communicative competence (e.g., Bachman, 1990; Canale & Swain, 1980) for many years and more recently in Celce-Murcia's (2008) formulation, which includes the use of speech acts under interactional competence. Although a variety of ways to assess pragmatic competence such as various types of discourse completion

tasks, role play, multiple-choice tasks, and the like (see Hudson & Brown, 2001; Hudson, Detmer, & Brown, 1992; Ishihara, 2009) can be applied, there are still issues such as authenticity, validity, reliability, and ease of administration. Furthermore, nonverbal behavior that accompanies speech acts is also an important aspect of pragmatic competence that should not be ignored (Jungheim, 2004, 2008, 2009b; Kendon, 1995, 2004; Yamashita, 1996).

Any written test of pragmatics will fail to tap learners' nonverbal ability (Jungheim, 2001). Gestures, head movement, and facial expression can be crucial to the performance of speech acts such as the refusal of an offer, invitation, or suggestion. Learners of Japanese as a second language, for example, may not be as good at understanding some gestures associated with refusals of an offer (Jungheim, 2006, 2008). They also gesture and use head movement slightly differently from native Japanese speakers (Jungheim, 2004), and the same can be said of some Japanese learners of English in relation to refusing (Gass & Houck, 1999). In a study of native Japanese speakers' perception of the optimal refusal of an offer (Jungheim, 2009a) in which negation, gestures, and facial expression were manipulated, participants judged persons performing refusals accompanied by a blank facial expression to have the strongest intention to refuse, regardless of whether or not the refusal included negation or a gesture. If nonverbal behavior is an important part of some speech acts, only a test of pragmatic competence that includes an actual physical performance, such as in role playing (Yamashita, 1996), will give learners the opportunity to show their true competence.

All things considered, a test should be evaluated for its usefulness (Bachman & Palmer, 1996), which includes reliability, construct validity, authenticity, interactiveness, impact, and practicality. Bachman and Palmer (1996) also note "In a classroom test...the teacher may want to utilize test tasks that will provide higher degrees of authenticity, interactiveness, and impact" (p. 19). The last three in particular are problematic for many of the methods for assessing pragmatic competence.

Collecting and evaluating actual performance data from a large number of learners is time-consuming and labor-intensive. In an attempt to deal with these problems, the purpose of this report is to describe the development of an Internet-based means of providing visual stimuli using video clips to collect video data from Japanese EFL learners performing a series of speech acts in authentic situations. Although rating may prove to be somewhat subjective unless clear rating guidelines are drawn up, this test, hereafter referred to as the e-Prag Test, will still have some advantages of a personal-response type test and may help overcome disadvantages such as planning and organization difficulties and difficulties for students, which have been pointed out in previous research (Brown, Hudson, Norris, & Bonk, 2002). This system can provide flexibility for the teacher/researcher and learners by enabling test administration to be carried out on-demand with video data stored on a server allowing teachers/researchers to carefully view and rate learners' performances at their leisure.

The First Stage of Development

In the first stage of development it is important to decide what aspects of pragmatic competence will be tested and to create contexts for their performance. This test is intended for Japanese university

students, so in order to enhance validity, the situations need to be authentic ones that a Japanese English student could be expected to encounter even within a Japanese university. Accordingly, the overall theme of this test is speech acts for university life. The range of speech acts covered in this test may seem a bit narrow considering the broader context of university life. However, as a classroom testing instrument, it is intended to provide a range of speech acts that might be introduced in a semester English class, for example, rather than a larger number that is further complicated by power, social distance, and imposition.

Pilot Instrument

Participants

Participants in the first stage pilot testing included three native speakers (one male, two females) and 11 Japanese students (five females, six males). They were recruited from among the researcher's students, their acquaintances, and his colleagues. Since the test uses visual stimuli, a range of virtual interlocutors was chosen to appear in the videos. They included males and females, younger and older persons, and different nationalities (British, Australian, American, and "other").

Instrument

This test is intended to be a kind of performance test simulating real situations that the students could encounter on campus and that would require English to respond appropriately. Specific speech acts were chosen on the basis of the researcher's experience and in consultation with a Japanese colleague who specializes in teaching and testing pragmatics. For example, this researcher often receives requests from students framed with "I want you to..." Only recently a student stopped me in the hallway and said, "I want you to check my English." This is an example of crosslinguistic influence from the Japanese *shite hoshii* (I want you to...). Feedback from Japanese learners of English indicates that they are clearly not aware that this is an inappropriate way to make a request to a teacher. A total of five speech acts were chosen for this test: an apology, responding to a complement, a request, a suggestion, and a refusal.

The following is the list of the situations chosen to elicit the speech acts:

1. Apologizing to a foreign student for knocking over his guitar case
2. Responding to a complement from a foreign student about a new sweater
3. Requesting a foreign professor's help with a report
4. Suggesting how a foreign student could buy a cheap train ticket
5. Refusing a foreign teacher's offer to help with a report because the student had a part-time job.

There were also different degrees of power, social distance, and imposition included in the situations.

Videos for the pilot stage of this study were taken at various locations on a university campus,

including a hallway, a professor's office, a location outside a classroom building, and a classroom, in order to enhance the authenticity of their content for this test. An Australian research student, an America professor, a British professor, and a Japanese student playing the part of a foreign student appeared in the videos.

A role-play format was adopted since this is a common means for practicing as well as testing the acquisition of speech acts and also because role plays "have the advantage over authentic conversation that they are replicable...." (Kasper & Dahl, 1991, p. 20), an important quality for a classroom-based test of interlanguage pragmatics. Although there may be some question about the authenticity of this semi-role-play format, Kasper and Rose (2002) suggest that role plays can be used because "inauthentic" does not necessarily mean "invalid" (p. 80). In the case of this test the format has the advantage of assuring that all test takers receive the same input, something that is uncontrollable in truly authentic data.

<p>E-Prag Testing</p> <p>Baseline Elicitation V. 1.8</p>	<p>University Life</p> <ul style="list-style-type: none"> You are going to watch a series of five video clips of various persons at a university. Imagine that you are a young student and answer each person as quickly as possible <u>in English</u>. 	<p>Scenario 1</p> <ul style="list-style-type: none"> You have just come back from a trip abroad, and your Australian friend comments about the sweater that you bought there. What will you say to her?
	<p>Scenario 2</p> <ul style="list-style-type: none"> You are a student who visits a teacher's office for help with your report. What will you say to him? 	
<p>Scenario 3</p> <ul style="list-style-type: none"> You are a student who knocks over a foreign student's guitar case while walking across the campus. What will you say to him when he looks at you? 		<p>Scenario 4</p> <ul style="list-style-type: none"> Your friend wants to take a trip to Kyoto and would like to buy a cheap train ticket. What will you say to her?
	<p>Scenario 5</p> <ul style="list-style-type: none"> Another professor wants you to come to his office to help you with your report, but you have no time today because of your part-time job. What will you say to him? 	
<p>Thank you for your cooperation!</p> <p>Do you have any comments about the situations or videos?</p>		

Figure 1. PowerPoint slide show used to pilot the videos and scenarios

Videos were subsequently arranged in a PowerPoint presentation in order to elicit baseline data from native English speakers and Japanese learners of English and to see if there were any problems with the instructions or the administration of the test as a whole that needed to be revised. Figure 1 shows the slides containing the instructions and videos clips used to elicit baseline data and pilot the video stimuli.

Procedure

The pilot version of the materials was administered in the researcher's lab. Participants were seated in front of a notebook computer, and their responses were recorded for subsequent analysis with a simple Sony Cyber-shot digital DSC-T700 camera. Using such a camera had an advantage over many video cameras in that it was easy to just insert the memory card into a computer's slot and open the files in a movie editor, in this case Adobe Premiere Pro CS4. A blue background was installed behind the participants, so that the videos of their performances could be superimposed over various scenes for use in the e-Prag Test as sample answers, as will be described below under Second Stage Development. The researcher operated the computer by remote control progressing through the slides when the participants indicated that they were ready to proceed. At the end of the test participants were given the opportunity to comment about their experience and note any problems that they might have had responding to the video clips.

Responses were transcribed and entered into a spreadsheet with additional coding for any gestures, head movement, and facial expression that accompanied the performance of the speech acts.

Pilot Results

The raw data from the participants' responses on the pilot test can be found in the Appendix. With only a few exceptions, learners were able to respond to the video stimuli with little difficulty. Since the actors remained on the screen after the videos finished, participants had a virtual interlocutor for their responses. As in previous research (Jungheim, 2006, 2008), many participants responded to the virtual interlocutors as they might have to a real person in a role play by including gestures and other nonverbal behaviors. For example, American participant NS 003, for example, tilted her head in a hedging move with her refusal. This kind of head tilt is used in English and Japanese alike to soften a refusal (Jungheim, 2004). In another example, Japanese participant NNS 006 gestured as if praying when apologizing for knocking over a student's guitar case. Because these contained salient nonverbal behaviors, they were chosen to be included in the e-Prag test as sample responses.

A number of problems did arise that needed to be resolved. In the version administered to the first three participants in the piloting, the scenario eliciting an apology for knocking over the guitar case appeared first. Since the actor did not say anything but only stared angrily, the first participants did not know when they were supposed to answer. This problem was solved for the pilot test by

moving the item from the item one position to item three and beginning with the item for responding to a complement. Also, the lack of an example item would have helped test takers to better understand how to respond, so an additional scenario would be needed for the e-Prag Test's example item.

There were also a few cases where nonnative speaker participants were not able to respond in some scenarios. In addition, although most of the participants had a reasonable level of English proficiency, NNS 007 and NNS 011 seemed to have a little more trouble than others, which may indicate a threshold for the overall proficiency needed to take this test.

Another problem that the researcher observed during the piloting concerned the instructions appearing before each video. Since the instructions all ended with the expression "What will you say...?" some participants hesitated and thought that they needed to respond after reading the instructions, until they saw the video in the next slide. Confusion caused by separating the instructions and the video into separate screens required attention for the construction of the online version of the e-Prag Test. The first-stage pilot version also permitted test takers to comment about the test at the end, so it would be useful to include such a function in the second stage of development. This could contribute to the improvement of the test as well as the quality of instruction in the event that the test is used in conjunction with actual classroom language learning.

The Second Stage of Development

In the first stage of development the researcher chose the overall theme of the test as speech acts for university life to enhance its validity for the target population of Japanese university English language learners. Five speech acts were chosen as ones that students could be expected to have an opportunity to use at a large Japanese university such as the researcher's institution where there are many foreign professors and students. Furthermore, appropriate situations and interlocutors were designated as ones that such students could encounter on campus during their student years. The researcher subsequently took videos of native speakers eliciting the target speech acts with care to make the scenes as realistic as possible. Then the videos and instructions were arranged in a slide show to construct a basic pilot instrument. Subsequent pilot test administration, analysis of the videos of the test takers, and review of input from the participants in the pilot study indicated issues that needed to be addressed in the construction of the first version of the e-Prag Test. The purpose of this section is to describe how these issues were addressed and how the online version of the test was created.

e-Prag Test Construction

The results of the piloting indicated that, although there were some problems, they could be solved, and all of the videos and scenarios could be used without revision. The first step then was to create a website to host the e-Prag Test. Fortunately there was enough space available on the university server to host the test on the researcher's own website. The next step was to revise

instructions appropriately for the online version. Additional videos were then created for an example item with a sample answer. This was followed by the superimposing of pilot video responses over appropriate scenes in order to use them as sample answers. Finally, a form was created for test takers to input personal information and comments or questions.

All the web programming was done by the researcher using Adobe Dreamweaver CS4 for the pages themselves and for the conversion of videos to Flash Video (flv) files that could be embedded and played back directly on the website at the test takers' convenience. Video editing was done with Adobe Premiere Pro CS4, which made it possible to superimpose the pilot test participants' videos over scenes that matched the stimulus videos.

Table 1. Comparison of Instructions of Pilot and e-Prag Test Versions

Item	Video Pilot Version	e-Prag Test Version
Intro	You are going to watch a series of five video clips of various persons at a university. Imagine that you are a young student and answer each person as quickly as possible in English.	You are going to watch a series of six short videos of situations involving English speakers you might interact with during your university years. Imagine that you are a student at the university in the videos and answer the person you see in each one as quickly as possible in English after watching it. You will have a chance to see an example response for each situation and can try to do it again, if you like. Now turn on your webcam to record your answers.
0		You are visiting the house of a friend who you have not seen for a while. Your friend greets you at the door. Play the video and then respond to your friend.
1	You have just come back from a trip abroad, and your Australian friend comments about the sweater that you bought there. What will you say to her?	You have just come back from a trip abroad, and your Australian friend comments about the sweater that you bought there. Play the video and then respond to your friend.
2	You are a student who visits a teacher's office for help with your report. What will you say to him?	You are a student who visits a teacher's office for help with your report. Play the video and then respond to your teacher.
3	You are a student who knocks over a foreign student's guitar case while walking across the campus. What will you say to him when he looks at you?	You are a student who knocks over a foreign student's guitar case while walking across the campus. Play the video and then respond to him when he looks at you.
4	Your friend wants to take a trip to Kyoto and would like to buy a cheap train ticket. What will you say to her?	Your friend wants to take a trip to Kyoto and would like to buy a cheap train ticket. Play the video and then respond to her question.
5	Another professor wants you to come to his office to help you with your report, but you have no time today because of your part-time job. What will you say to him?	Another professor wants you to come to his office to help you with your report, but you have no time today because of your part-time job. Play the video and then respond to his suggestion.

*This scenario was number 0 for the first three participants, but the order was changed because they did not understand when to respond.

In the interest of test security, the top page was programmed to be password-protected. Furthermore, instructions were revised slightly to make them more appropriate for the online version of the e-Prag Test as shown in Table 1. Instructions were inserted on the same page as the related video clip for each scenario to facilitate the operability of the test and to help clarify the method of response. A screen capture of the sample item was used to graphically illustrate how to navigate the test, including instructions for how to go to the sample item as explained on the introduction page of the test. The example item, greetings between friends, was intentionally not a speech act, because the researcher wanted to avoid any input that might influence later answers on the test. This video was taken in the entrance hall of the researcher's house using his daughter and her friend who visits after being away for a while.

The inclusion of sample responses and the option of repeating each item was intended to provide immediate feedback and a positive washback effect, defined as "the effect of testing on teaching and learning" (Hughes, 1989, p. 1). Thus, learners could compare their own answers and revise them if they felt that the sample answer input was useful. Sample answers were not considered to be definitive answers but merely examples of possible responses, since authentic conversation would allow for a myriad of responses. For this reason, sample answers also included responses of nonnative speakers, one of which had a nonnative speaker self-correcting a grammatical error, which gave it a more "authentic" feeling. The researcher's intention was to allow for reconsidered responses that could then be rated in the same way as the test takers' initial responses. Any improvement in the performance of the speech acts could be attributed to the positive washback effect of the sample answers. Thus, the sample answers served, in a sense, as a kind of consciousness-raising input. Table 2 shows the stimulus and sample answer for each item, including the example item.

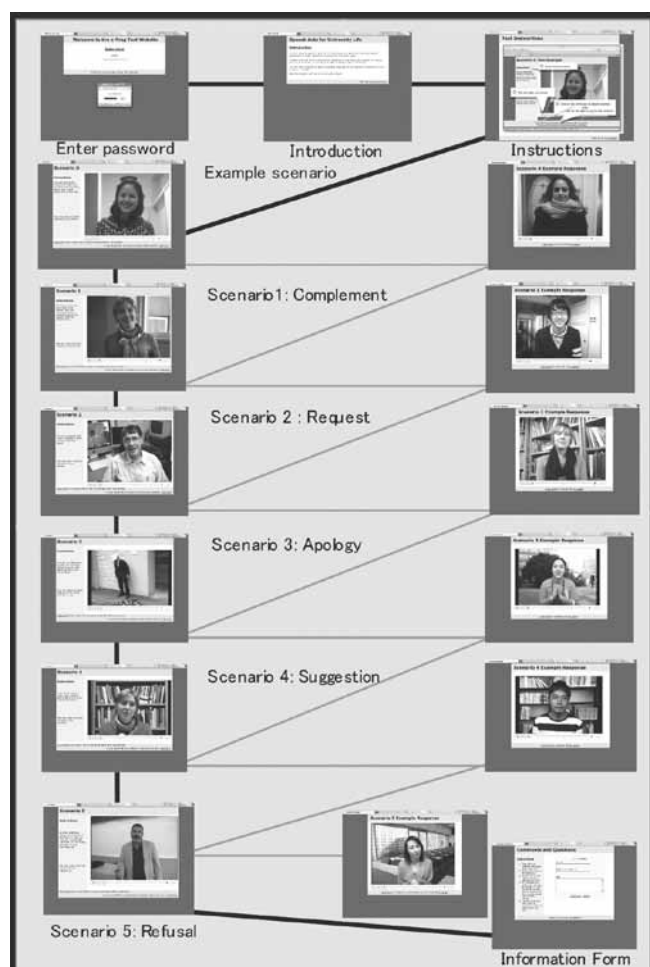
Table 2. Stimuli for each item and related sample answers

Item	Stimulus	Sample Answer
0	Hey! It's been a long time. How're you doing?	Not bad. How about you?
1	Hi. Welcome back. Did you get that sweater on your trip? I really like it.	Thank you.
2	(knock) Yes? How can I help you?	Um, I'm having a bit of trouble with the report that you set, and I was wondering, if you have time, whether you could maybe go through it with me?
3	(Guitar knocked over; student glares angrily)	Sorry! I'm very sorry. I didn't meant that. I didn't mean it.
4	I want to go to Kyoto. Do you know how to get a cheap ticket?	I think you should buy on the internet. You can find a cheap ticket.
5	Why don't you come and talk with me if you're having trouble with that research paper?	Thanks for your help but I have a part-time job today. Do you have time maybe, perhaps tomorrow?

Since test takers will be taking the test remotely, it is important to provide some way to give and get feedback. This was accomplished by programming a form with instructions that could upload the test takers' personal information and any comments or questions related to the e-Prag Test. The instructions are:

1. Turn off your webcam and save your video file.
2. Write your name in the title box.
3. Write your e-mail address in the appropriate box.
4. In the large box write your (a) age, (b) length of time studying English, and (c) amount of time spent abroad and country names.
5. Finally, write any questions or comments that you have about the test.

The additional information allows for a more detailed analysis of the test data, a more finely tuned interpretation of individual participants' responses in light of their experience, and input that could be applied to the possible revision of materials used in the classroom.



Note. The black lines indicate the order of the required steps. The grey lines indicate the optional viewing of sample answers on the left and the possibility to return and repeat an item.

Figure 2. Screen captures of e-Prag Test showing the overall flow of the test.

Figure 2 shows screen captures of each screen of the e-Prag test arranged in a flow chart. Black lines indicate the order of the required steps in the test, and grey lines indicate the optional viewing of sample answers and the possibility to return and retake any item. Since videos are taken of each test taker, it will be possible to see if anyone skips an item the first time around and views the sample answer before going back and answering. This would result in a rating of zero for not doing the item first. Note how the superimposed videos in the sample responses appear to have really been taken in the same settings as the video stimuli.

Preliminary Analysis of the e-Prag Test

Bachman and Palmer (1996) suggested six qualities that a test should have to satisfy its usefulness and justify using the scores for decision-making. These are reliability, construct validity, authenticity, interactiveness, impact, and practicality. This section will examine how the e-Prag Test fulfills these criteria and also consider their relationship to what Chapelle (2001) identifies as positive and negative attributes of technology in computer-aided language testing (CALT).

First and foremost is reliability, or test consistency. In the case of the e-Prag test and its small number of items, it is essential that very clear criteria be created for rating the responses in order to minimize inconsistencies among test scores. Although the pilot study responses may be useful for creating rating scales, and the pilot test video responses could even be used to train independent raters, it is necessary to further collect data in the form of the online version of the e-Prag test to obtain a set of responses from native speakers and language learners under the actual conditions in which the test was intended to be administered. As this is a report of the test development only, the researcher has not yet entered the third stage of development, the validation stage. Therefore, reliability is an aspect that remains to be examined in the next stage. Since scores will be given by raters based on a scale, this test has the positive attribute of partial-credit scoring that may be more precise and result in a greater variance of scores than dichotomously-scored items, e.g. multiple choice (Chapelle, 2001), while having no particular negative attributes as CALT in relation to reliability.

The next most important quality of a test is construct validity. To what extent can we justify our interpretation of scores and generalize them to the target language use (TLU) domain? As an ongoing process (Bachman & Palmer, 1996), we are clearly not ready to justify this at the moment, although we have addressed the issue of validity at the initial stages by choosing a particular context for the test and carefully deciding speech acts that test takers could be expected to need in that context. A positive attribute of this test in relation to construct validity is related to it consisting of open-ended items, which "are less likely...to be affected by systematic test-taking strategies" (Chapelle, 2001, p. 115).

As for authenticity, once again the choice of scenarios and the presentation of the video stimuli in "authentic" contexts on the university campus make this a positive attribute because they do simulate actual conditions as closely as possible. The test takers respond spontaneously to oral and visual input

as they might in real life. The main drawback is that they do not receive immediate feedback from an interlocutor as to the understandability or appropriateness of their responses. The sample responses partially compensate for this by providing feedback, albeit in a less-than-authentic form. This lack of complete authenticity can be cited as a negative attribute. There do not seem to be any other negative attributes to this test for authenticity of content, since it is closely related to the actual performance of the speech acts, but with an advantage over face-to-face interaction of having control over the stimuli.

The e-Prag test has positive attributes for interactivity, because the test requires the test takers to relate their language knowledge of speech acts to the individual scenarios. The test also engages their topical knowledge and affective schemata by presenting stimuli in the form of familiar situations that they could encounter in everyday life on the university campus. The variation of power, social distance, and imposition in the scenarios may also engage their affective schemata.

This test satisfies impact in a number of ways, in particular the impact on learning and teaching through the washback effect as described above. It could also have an impact on the classroom by encouraging teachers and learners to create their own video materials. Digital cameras that can take videos of sufficient quality are relatively inexpensive and most computers come with video editing software such as Moviemaker or iMovie, which open the world of moviemaking to most students in a developed country like Japan. A negative attribute of this test is the expense of programming that limits test construction to those who can afford to do it or those who have programming expertise. Although it might not result in extra expense for individual learners, it could be economically unfeasible for institutions.

Finally, the test has a positive attribute for practicality, since it frees learners and teachers to play their parts in the testing process at their own convenience. "Internet-delivered tests add flexibility of time and place for test delivery" (Chapelle, 2001, p. 115). On the other hand, as Chapelle points out a negative attribute of this kind of test is test security. Once someone takes the test, its content can be easily communicated to other test takers. This can only be dealt with by creating alternate forms, something that could be costly and time-consuming.

On the whole, the e-Prag test has more positive than negative attributes. Unfortunately, further analysis is required in the third stage of development to fully understand the extent to which the test satisfies the seven qualities of a good test. The following section briefly describes the third stage of development.

The Third Stage of Development

As it stands, the e-Prag test in its present form is basically a beta version that needs to be thoroughly tested and revised. The third stage of development will involve further piloting and revision, additional programming to permit uploading of videos directly from the test-takers' webcams from either their homes or from computers on campus, and programming of a more efficient feedback form.

Preparation for larger-scale piloting of the beta version of the test will be carried out without the uploading of webcam videos. In principle this will be done in the researcher's office, but it will also be carried out with participants who can record and save their video into a file either with a digital camera or a webcam. Once a sufficient number of participants have completed the test, detailed rating scales will be created based on grammaticality, the performance of the speech act itself, and the appropriateness of the speech act for the interlocutor, including the consideration of different degrees of power, social distance, and imposition. Although the composition of rating scales was considered at all stages of development, more data collected using these visual stimuli is needed before detailed scales can be constructed that are not plagued with problems such as inference (difficulty understanding what the scores infer), difficulties of assigning levels (need to satisfy multiple criteria at the same time), and differential weighting of components (hidden differentials in raters' weightings) (Bachman & Palmer, 1996). Additional video samples are also needed to create training videos to help raters calibrate their ratings.

Validity will be further addressed by correlating the results of the e-Prag Test with the results of other tests of pragmatic competence (see examples in Hudson, et al., 1992; Yamashita, 1996, 2001).

Conclusion

This report outlined the basic stages of the development of an online test of pragmatic competence called the e-Prag Test. Although there are many positive aspects of this test in its present form, a great deal of work remains to be done before it can be used by teachers and learners or by researchers.

The existence of problems that still need to be addressed can be attributed to the short time of one year that the initial project was allotted, as well as the limited funds that precluded the hiring of someone with the professional expertise to do the substantial amount of programming needed for a more sophisticated version. In addition, a system that would permit the uploading of videos from test takers' webcams would require a dedicated server that, unlike the university server, would permit greater flexibility regarding what can be uploaded and the use of more appropriate forms. The form used for feedback on this test is the only one that the university server allows. The researcher only has limited expertise in web programming.

One more issue that came up when colleagues viewed the beta version of the test was related to privacy. By doing the test in front of their own computers at home, there is the potential for the invasion of test takers' privacy, unless they make sure that a blank wall is behind them when they take the test. Of course, there is also the question of whether or not learners will accept having their videos taken. The privacy issue also includes measures that will be taken to delete or preserve the video files, as well as the need to inform test takers as to how this will be carried out. Security for this type of test is a complex issue that includes ethical considerations.

On the positive side, this project succeeded in developing a workable online test. Only through further piloting of the test can it be revised in a way that will make it possible to administer in a real testing situation. In the final analysis, this requires funds and time that may not be available. An easy-to-administer online test of pragmatic competence that would give test takers the opportunity to respond orally in an open-ended test format, and include the possibility of evaluating nonverbal behavior accompanying speech, would be a useful contribution to language learning and language testing.

References

- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in Practice*. Oxford: Oxford University Press.
- Brown, J. D., Hudson, T., Norris, J., & Bonk, W. J. (2002). *An investigation of second language task-based performance assessments* (Technical Report #24). Honolulu: University of Hawai'i, Second Language Teaching & Curriculum Center.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second-language teaching. *Applied Linguistics*, 8, 67-84.
- Celce-Murcia, D. (2008). Rethinking the role of communicative competence in language teaching. In E. A. Soler & M. P. S. Jordà (Eds.), *Intercultural language use and language learning* (pp. 41-58). Dordrecht, The Netherlands: Springer.
- Chapelle, C. A. (2001). *Computer applications in second language acquisition: Foundations for teaching, testing and research*. Cambridge: Cambridge University Press.
- Gass, S. M., & Houck, N. N. (1999). *Interlanguage refusals: a cross-cultural study of Japanese-English*. Berlin: Mouton de Gruyter.
- Hudson, T., & Brown, J. D. (Eds.). (2001). *A focus on language test development: Expanding the language proficiency construct across a variety of tests* (Technical Report #21). Honolulu: University of Hawai'i Second Language Teaching and Curriculum Center.
- Hudson, T., Detmer, E., & Brown, J. D. (1992). *A framework for testing cross-cultural pragmatics* (Technical Report #2). Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Ishihara, N. (2009). Teacher-based assessment for foreign language pragmatics. *TESOL Quarterly*, 43(3), 445-470.
- Jungheim, N. O. (2001). The unspoken element of communicative competence: Evaluating language learners' nonverbal behavior. In T. Hudson & J. D. Brown (Eds.), *A focus on language test development: Expanding the language proficiency construct across a variety of tests* (pp. 1-34). Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.
- Jungheim, N. O. (2004). Hand in hand: A comparison of gestures accompanying Japanese native speaker and JSL learner refusals. *JALT Journal*, 26(2), 127-146.
- Jungheim, N. O. (2006). Learner and native speaker perspectives on a culturally-specific Japanese refusal gesture. *International Review of Applied Linguistics*, 44(2), 125-143.

- Jungheim, N. O. (2008). Language learner and native speaker perceptions of Japanese refusal gestures portrayed in video. In S. G. McCafferty & G. Stam (Eds.), *Gesture: Second language acquisition and classroom research* (pp. 157-182). New York: Routledge.
- Jungheim, N. O. (2009a, March). *In search of an optimal Japanese Refusal: The use of language, gesture, and facial Expression to clarify intention* Paper presented at the Conference of the American Association for Applied Linguistics, Denver, CO.
- Jungheim, N. O. (2009b). Japanese refusals and obligatory contexts for gestures. *Eibungaku [English Literature]*, 95, 1-18.
- Kasper, G., & Dahl, M. (1991). *Research methods in interlanguage pragmatics* (Technical Report #1). Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.
- Kasper, G., & Rose, K. R. (2002). *Pragmatic development in a second language*. Malden, MA: Blackwell.
- Kendon, A. (1995). Gestures as Illocutionary and discourse structure markers in Southern Italian conversation. *Journal of Pragmatics*, 23, 247-279.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Yamashita, S. (1996). *Six measures of JSL pragmatics* (Technical Report #14). Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.
- Yamashita, S. (2001). Using pictures for research in pragmatics: Eliciting pragmatic strategies by picture response tests. In T. Hudson & J. D. Brown (Eds.), *A focus on language test development: Expanding the language proficiency construct across a variety of tests* (pp. 35-56). Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.

Appendix

Pilot Video Responses

Participant	Apology	Respond to a Complement	Request for Help	Suggestion	Refusal
NNS 001		Sounds good. I'm glad to hear that.	Would you mind to seeing my report?	Yes, I know the place you can buy a cheap ticket.	I'm sorry, but I have to go do my part-time job.
NNS 002	I'm sorry.	Oh, yes, it looks great.	Um, excuse me; I want you to check my homework.	Um, I'm sorry I don't know quite pretty sure, but you can ask to, ask in station.	I'm sorry, but today I have part-time job. If you have some time, the other day I will go.
NNS 003	I'm sorry.	Your sweater very suits you.	Will you talk about my report.	Ah, sorry. I don't know in detail, so please ask station staff.	Sorry. today I'll go to my part-time job so how about another day?
NNS 004	I'm sorry.	Thank you.	I would like you to tell me the, tell me how to, how to write this report.	You had better to buy a <i>seishunjubachi kippu</i> .	I'm sorry. today I have part-time job, so do you have time, do you have time, ah tomorrow?

NNS 005	I'm sorry.	Thank you.	I have a problem with my report, so I want to, I want you to give me a, some advice.	I think you should buy on the Internet. You can find a cheap ticket.	I'm sorry. Today I'm busy, because I have a part-time job, so I can help you another day.
NNS 006	Sorry! I'm very sorry. I didn't meant that. I didn't mean it.	Oh, thank you. This is the sweater I bought there, so I like this color.	Professor, sorry, ah, I wonder, you have some time to, you have some time to help me, because I have some trouble with my report. So I, would like to, so, so if you are convenient, I hope you can help me.	Ah ok. I'll help you search on the Internet. I'll try it.	Oh, thank you, Professor. Thank you for your kindness. But today I have, I don't have time, because of my part-time job, so would you, ah, can I, could we, could we meet on another time?
NNS 007	Don't put it on, <i>chigau</i> . You are, you are wrong.	Yes. I like it.	I want you to check my report.	Try to buy a ticket for kids.	
NNS 008	I'm sorry.		I want to check my, ah, I want you to check my report.	You can, you can, ah, <i>nandaro</i> . Ah, you can buy cheap ticket on Internet.	I want you, ah, check my report another time.
NNS 009	I'm sorry.		Please give me, ah, advice about my thesis.	I don't know.	I'm sorry I have, ah, part-time job today.
NNS 010	I'm so sorry.	Thank you.	I want my report checked.	You should go to the station and ask someone.	I'm sorry, but I have a part-time job today.
NNS 011	I think he should say, "I'm sorry."	It was cold in Australia.	I lost my USB.	Yes, I know. <i>Seishunjuhachi kippu</i> is very cheap.	I'm sorry. I have a part-time job tomorrow. I will be free, ah, today, I will be free tomorrow.
NS 001	Oh, I'm sorry. I'm sorry. Excuse me.	Oh, thanks very much. Yeah, I got it on the trip.	Hi. Um, do you have a moment? I need some help with my report.	I think the best way is to go to the travel office at one of the main stations like Tokyo station or Shinagawa station and ask there. They speak good English, so they'd be able to help you.	I'm sorry. I will do that, but today I have to go to my part-time job. Sorry about that.

NS 002	Oh I'm sorry. Is it ok?	Yeah, I bought it in Sydney. It's pretty nice huh.	Um, I'm having a bit of trouble with the report that you set and I was wondering, if you have time, whether you could maybe go through it with me?	Yeah, I think I know of a few websites. Ah, let me have a look into it, and I'll get back to you.	Um, well, I, that's really kind of you, and I would like to come today, but unfortunately I have to work, so do you think, um, we could do it tomorrow, if you have time?
NS 003	Sorry. Excuse me.	Thanks. Yeah, I got it in Sydney.	Hi. I was wondering if you could look over a report for me.	Ah, I would go to the JR station and ask for their help. It's the green sitting man.	Thanks for your help, but I have a part-time job today. Do you have time maybe, perhaps tomorrow?
