Title: An inquiry into pragmatic data collection methods

Author: Yuh-Fang Chang

Date: 2006/6

Abstract

While the significance of validation of data collection instruments in speech act research has been recognized and has attracted considerable interest, most validation studies employed a between-subjects design. In so doing, it is possible that the differences were caused by the group effects rather than different data collection techniques. This study attempts to resolve the above-mentioned problems by employing a within-subject design and comparing refusal responses obtained from the written DCT, closed role play (oral DCT), and open role play in terms of response length, frequency of semantic formulas and the combination pattern of the semantic formulas. Sixty female college students participated in the study. It was found that the three methods elicit similar samples in terms of response length, mean number of semantic formulas, frequency of semantic formulas and the combination pattern of the semantic formula.

Keywords: data collection method, DCT, role play, refusals

Introduction

In contrastive pragmatics and interlanguage pragmatics research, the most frequently used instrument for eliciting speech act data is the discourse completion test/task (DCT) (Bardovi-Harlig & Hartford, 1993; Rose & Ono, 1995). In spite of its popularity, the validity of this method for eliciting speech act data has been questioned. Several researchers have reported the discrepancy between DCT and spoken data (e.g., Beebe and Cummings, 1985, 1996; Golato, 2003; Rintell & Mitchell, 1989; Sasaki, 1998; Turnbull, 2001). Other validation studies compared data produced in responses to DCTs and multiple choice questionnaires and found significant differences between the two (e.g., Rose, 1994; Rose & Ono, 1995). In addition, research also documented that the prompt provided in the DCTs affects the subjects' choice of strategies (e.g., Bardovi-Harlig & Hartford, 1993; Billmyer & Varghese, 2000; Blum-Kulka et al. 1989; Johnston, Kasper & Ross, 1998; Rose, 1992).

While the significance of validation of data collection instruments in speech act research has been recognized and has attracted considerable interest, the research findings, however, are not unequivocal. For example, while Rintell and Mitchell (1989) found little difference in data elicited from oral versus written versions of the discourse completion test (DCT), Yuan (2001) showed that the oral DCT generates a significantly larger number of natural speech features than the written DCT.

Aside from the inconclusive findings, previous studies also suggested that special care in the research design of further study is needed. Most validation studies employed a between-subjects design. That is, each research participant took part in only one condition of the study. In so doing, it is possible that the differences were caused by the group effects rather than different data collection techniques. To avoid the possibility of group effects, a within-subjects design (i.e., using the same

participants for all conditions of the study) would be more appropriate for a methodological validation study. Alderson et al. (1995) and Bachman (1990) also pointed out the significance of employing a within-subjects design when checking the validity and reliability of different methods as evaluation measures.

Moreover, few validation studies collected refusal data (see Table 1). Most of the studies examined the methodological issues through request data. The speech act of refusals differs from other speech acts in that it functions as a response to another act rather than an act initiated by the speaker. Because of this, the range of the possible refusal responses is broader and the time allowed to plan for the act is more limited than the speech acts of requests and compliments. Hence, whether the research findings obtained from requests or compliments hold up across the speech act of refusals requires further investigation.

Furthermore, the possible effects of proficiency level on the participants' responses to different methods have remained unexplored. Variability has been one of the characteristics of interlanguage. It is very likely that the learner's stage of development will have a great impact on whether and how instrument-based pragmatic variability is exhibited. This study is motivated by the limitations of previous research and designed to answer the following research questions:

(1) Do L2 learners' refusal responses differ according to different data collection techniques? If so, how?
(2) What are the effects of proficiency levels on the participants' responses to different methods?

Literature Review

This section reviews studies concerning the data elicitation techniques in the area of speech act research: discourse completion tests, observation of naturally occurring speech, and role play.

A. <u>Discourse completion tests</u>

The discourse completion test/task is a written questionnaire in which a situation is given specifying the setting, the social distance between the participants, and their respective social status, followed by a brief dialogue with a blank for the respondent to provide a written response. The subjects are asked to write down what they think they would say in the described situations. This type is called open-ended elicitations. Another type of DCT provides a follow-up response after the blank which helps to cue the respondents as to the appropriate content of the response to be filled in. An example of the second type of DCT is:

> You are the owner of a bookstore. One of your best workers asks to speak to you in private.
> Worker: As you know, I've been here just a little over a year now, and I know you've been pleased with my work. I really enjoy working here, but to be quite honest, I really need an increase in pay.
> You:_____
> Worker: Then I guess I'll have to look for another job.

The discourse completion task was originally employed by Blum-Kulka (1982) and has been an extensively used elicitation method in cross-cultural speech act studies. Beebe & Cummings (1985), in their study of data elicitation methods, claimed that the DCT is a highly effective means of:

(a) gathering a large amount of data quickly;
(b) creating an initial classification of semantic formulas and strategies that will occur in natural speech;
(c) studying the stereotypical, perceived requirements for socially appropriate (though not always polite) responses;
(d) gaining insight into social and psychological factors that are likely to affect speech and performance; and
(e) ascertaining the canonical shape of refusals, apologies, partings, etc., in the minds of the speaker of that language (p.13).

Furthermore, because this elicitation technique allows the researcher to control for situation, the researcher can manipulate the variables of interest (e.g., gender, social

distance, age).

However, because the DCT elicits written responses, certain kinds of information such as elaborated responses typically found in naturally occurring interactions, prosodic, and nonverbal features of oral interaction cannot be obtained through this data collection method (Cohen, 1996; Hartford & Bardovi-Harlig, 1992). Wolfson, Marmor, and Jones (1989) pointed out that short decontextualized written responses may not be comparable to authentic spoken interaction. Beebe and Cummings (1985, 1996), likewise, noted that DCT responses do not adequately represent:

> (a) the actual wording used in real interaction;
> (b) the range of formulas and strategies used (some, like avoidance, tend to be left out);
> (c) the length of response or the number of turns it takes to fulfill the function;
> (d) the depth of emotion that in turn qualitatively affects the tone, content and form of linguistic performance;
> (e) the number of repetitions and elaborations that occur; or
> (f) the actual rate of occurrence of a speech act—e.g., whether or not someone would naturalistically refuse at all in a given situation (p.14).

B. <u>Ethnographic observation</u>

Due to the limitations of the discourse completion test/task, some researchers (Wolfson, 1981, 1986; Wolfson, Marmor, & Jones, 1989) stated that the best approach to elicit data which enable researchers to learn the social and linguistic constraints on a particular speech act is ethnographic observation. According to Watson-Gegeo (1988), ethnographic observation involves the observation of naturally occurring speech with exact recording about participants, the event, the location, and social settings.

The strengths of this data elicitation approach include: 1) the assurance of the

internal validity of the study as they represent spontaneous natural speech as it really is; and 2) the provision of rich contextual information which helps researchers to examine the pragmatic appropriateness of the speakers and reduces the likelihood of misinterpreting the situation.

However, several researchers have pointed out the limitations of ethnographic research (e.g., Bardovi-Harlig & Hartford, 1992; Kasper & Dahl, 1991; Rintell & Mitchell, 1989). First, it does not allow for researchers to control for contextual variables. Second, it is difficult to obtain a sufficient corpus of data for ethnographic observation because the occurrence of a particular speech act cannot be predicted. Third, the results obtained from the ethnographic approach cannot be replicated since it is very unlikely that the same situation will occur twice in real life. Fourth, Kasper & Dahl (1991) noted that it takes about 10 hours to transcribe a 1 hour audible tape in ordinary orthography. The ethnographic data entail transcription, therefore, the process is very time-consuming.

C. Role play

Since the naturally occurring data are difficult to collect, role plays are used to gather naturalistic data. There are two types of role play: closed role plays and open role plays. In both types, instructions which specify the roles, the initial situation and at least one participant's communicative goal are given to subjects. In closed role plays, subjects are asked to give a one-turn oral response.

On the other hand, open role plays often take several turns in discourse. Subjects are not instructed as to what conversational outcome needs to be reached or how such outcomes are reached. As Kasper and Dahl (1991) pointed out, open role plays "represent oral production, full operation of the turn-taking mechanism, impromptu planning decisions contingent on interlocutor input, and hence, negotiation of global and local goals, including negotiation of meaning, when required (p.228)."

The greatest strengths of open role plays are that they allow for researchers to examine a particular speech act behavior in its full discourse context, to observe how speech act performance is structured in sequence and how specific strategic choices affect interlocutor responses. Moreover, open role plays have a primary advantage over authentic conversation: they are replicable and enable researchers to compare the speech act performance between native speakers and non-native speakers (Kasper & Dahl, 1991, p.229).

One major disadvantage of role plays as a data collection technique is that they need transcribing and, thus, are time-consuming. In addition, subjects may exaggerate the pragmatic interaction in performing role plays, producing a speech behavior which would not have occurred in a real-life situation, since they know that the situation is not real life and their response won't threaten either the speaker's or the listener's positive or negative face. Larson-Freeman & Long (1991) also pointed out that subjects are apt to produce the item the researcher is interested in studying, which may threaten the validity of the study. Moreover, role-plays would most likely be audio- or videotaped, which may engender problems. It is very likely to make subjects feel uncomfortable about the taping and may affect their role play performance.

D. Studies comparing different types of production data

This section reviews the comparative studies. Several studies have been conducted with the purpose of assessing the validity of discourse completion test (DCT). Rose (1994) compared English NSs' and Japanese NSs' performance of requests in discourse completion tests and multiple choice questionnaires. The inspection of the data showed that although indirect request was the most frequent response to all situations, hinting or opting out occurred more often with the multiple choice questionnaires than discourse completion tests. Parallel findings were obtained

in a study by Rose and Ono (1995). As pointed out by Gass and Houck (1999), both studies challenge the validity of discourse completion tests, especially in a non-western context.

Several researchers have reported the discrepancy between DCT and spoken data. Yuan (2001) compared Kunming Chinese responses to compliments in written DCTs, oral DCTs, field notes and recorded conversations. Yuan (2001) found that the oral DCT generates a significantly larger number of natural speech features than the written DCT. Beebe & Cummings (1985, 1996) compared refusals performed by native speakers of English in response to real telephone requests and a single discourse completion item based on these requests. It was found that the telephone conversations were more elaborated than the discourse completion responses. In comparison with the discourse completion responses, the range of refusal strategies used in the telephone conversations appeared to be wider. However, the frequency and the content of refusal strategies used in both data types were similar.

Another study comparing discourse completion tests with naturally occurring data was conducted by Hartford and Bardovi-Harlig (1992). They compared native speakers' and non-native speakers' performances of refusals in academic advising sessions and in discourse completion tests. They also pointed out that data elicited from discourse completion tests show a more limited range of semantic formulas, and fewer status-preserving strategies.

The discrepancy, however, was not found by Rintell and Mitchell (1989). Rintell and Mitchell (1989) compared requests and apologies elicited by written and oral versions (also referred to as closed role plays) of the same discourse completion test, which were given to low advanced learners of English and to native English speakers. In contrast with the above-mentioned studies, the results showed little difference in data elicited from oral versus written versions of the discourse completion test (DCT).

They claimed that both elicitation methods provide data that are similar to spoken language rather than written language.

Unlike the aforementioned studies, which used a between-subject design, Sasaki (1998), employing a within-subject design, compared discourse completion tests and open role plays eliciting requests and refusals. Investigations of results revealed differences in response length and content. The findings, however, were based on only 12 participants, which calls into question their generalizability. Using a within-subject design, Yamashita (1996) conducted a multi-method comparison study. The data collection method comparison included: 1) a self-assessment, 2) an oral production test, 3) an open discourse completion test, 4) a role play, 5) a role play self-assessment, and 6) a multiple-choice discourse completion test. The results showed that all methods are reliable and valid except for the multiple-choice discourse completion test. In addition, the length of exposure to the L2 was found to correlate positively with two tests (oral production and role play).

Rather than contrasting DCT with other method, Dahl (1995) and Felix-Brasdefer (2007) compared authentic discussions with open-ended role plays. Substantial differences were found in the amount of talk and directness in the performance of face-threatening acts between two types of data. Dahl concluded that role plays are not representative of authentic interaction on these measures. Dahl, however, cautioned that the conclusion drawn from her relatively constrained type of role play may not be generalizable to more open and self-directed role play types.

As aforementioned, many validation studies employed a between-subjects design (see Table 1), which may introduce the group effect as a confounding variable in the study. In addition, few validation studies compared closed role play and open role play refusal data. Furthermore, the possible effects of proficiency level on the

participants' responses to different methods have remained unexplored. This study attempts to fill the gap by employing a within-subject design and comparing refusal responses obtained from the written DCT, closed role play (oral DCT), and open role play.

Methods

Subjects

Sixty female college students were selected to represent two different English proficiency levels (intermediate, advanced). Thirty intermediate level students were non-English major freshmen. Thirty advanced level students were English major seniors. To include only female subjects is to control for the gender differences.

Procedures

The purpose of this study was to compare learners' refusal responses from the written DCT, closed role play (oral DCT), and open role play. Two refusal situations were devised involving 2 stimulus types eliciting a refusal:1 request and 1 invitation (see Appendix A). The refuser's social status relative to the interlocutor involved 2 levels: equal and low. Participants' familiarity with the role in the situations was taken into consideration. Table 2 shows the classification of the discourse completion test/task.

To assure the authenticity of the language used, the participants were first instructed to do the closed role play, the open role play second and finally the written DCT. The interval between the tasks was two days. The rational for arranging the written DCT as the last task rather than the first one is to lessen the possibility for the written DCT responses to affect the oral responses. Since the difference between the closed role play and the open role play lies in that the former requires subjects to give a one-turn oral response whereas the latter requires subjects to take several turns in

discourse, the decision was made to arrange the closed role play as the first task and the open role play as the second one. Two-day interval between the tasks was arranged in order to decrease the practice effect. Instructions concerning how to do each task were provided by a research assistant to the participants before each task was carried out.

A. Closed role play (oral DCT)

In the oral DCT task, subjects were asked to listen to a tape-recording of situations identical to the situations in open role play and written DCT and then respond to them orally. To avoid the anxiety effect, two practice situations were added prior to the actual test situation, but discard in the data analysis. Subjects' responses were tape-recorded and subsequently transcribed.

B. Open role play

Open role plays were performed one at a time in a separate room (i.e., room A for situation 1; room B for situation 2). One teacher and one student were asked to be in each room to play the role of the interlocutor in each situation. Subjects' responses again were tape-recorded and subsequently transcribed.

C. Written discourse completion questionnaire/test

To do the DCT task, the subjects were instructed to write down what they think they would say in the described situations.

Data Analysis

The analysis of the data involved calculating and comparing the response length, mean number of semantic formulas, mean number of pause fillers, patterns of semantic formula combined in the refusal responses and reason/ excuse provided in the response fro each task. The response length was calculated by counting the total number of words in the refusal responses elicited from different methods. Pause fillers such as oh, um, uh and well were excluded from the word count.

The length of the refusal responses itself, however, did not provide much information on how the participants refused to each situation. In order to have a better understanding on what refusal strategies the participants used to refuse each situation, the refusal responses were further analyzed as consisting of a sequence of semantic formulas, based on the well-established taxonomy developed by Beebe et al. (1990). For example, if a respondent refused a request from his boss to stay late at the office, saying "I'm sorry. I can't stay this evening. I have a prior engagement. However, I will take care of it first thing in the morning," this was coded as: [expression of regret][negative ability][excuse][offer of alternative]. The mean number of semantic formulas, and the patterns of semantic formula combined in the refusal responses were examined. In addition, the semantic formula of "pause fillers" such as "oh", "um", "uh" and "well" were also calculated and compared since pause fillers have been reported to be a distinctive feature of spoken language. Since the open role play involved more than one turn, it is the refusal responses of the 1st turn that was used when compared with the other elicitation method. The 1st and 2nd turn of the responses in the open role play was also compared. The repeated measure ANOVA, and matched t-test were performed to detect whether the difference is significance.

Findings

The means and standard deviations for the response length of DCT, closed role play and open role play for each group are presented in Table 3. The results of ANOVA are displayed in Table 4. As seen in Table 4, there was a significant main effect for Proficiency for each situation (**situation1:** F=14.851, p<.0001; **situation2:** F=27.152, p<.0001). In general, the refusal responses produced by advanced students were significantly longer than that by the intermediate level students. There was no significant main effect for Task. Overall, in each situation, subjects responded at similar length across three elicitation methods. Table 5 displays the difference in the

means and standard deviations of the response length between the 1st turn and 2nd turn of the open role play. The participants overall had a longer response in the 1st turn than in the 2nd turn of the open role play task. The difference, however, was deemed significant for the situation 2. Table 6 displays the mean number of semantic formula employed in DCT, closed role play and 1st turn of the open role play. The results of ANOVA are presented in Table 7. No significant main effect was found for either task or proficiency on the mean number of semantic formula. In other words, the analysis revealed that neither of group performed differently across three elicitation methods. The means and standard deviations for the number of pause filler used in DCT, closed role play and open role play for each group are shown in Table 8. The results of ANOVA are displayed in Table 9. Table 9 shows that participants used significantly more pause fillers in role play tasks than in the DCT for both situations (**Situation1:** $F = 5.38$, $p=0.05$; **Situation2:** $F = 9.78$, $p<0.05$). Tables 10 and 11 illustrate quantitative comparison of the number of combination patterns of semantic formulas and the most common combination patterns across tasks. There were a total of 29 combination patterns of semantic formulas in DCT, 24 in the closed role play and 21 in the open role play for situation 1. As in situation 1, DCT yielded more combination patterns of semantic formula than the other two tasks in situation 2. The most common combination patterns included: [regret]-[negative ability]-[reason], [regret]-[reason]-[alternative], [regret]-[reason].

Table 12 presents the examples of the refusal responses for DCT, closed role play and open role play for each group. As seen in Table 10, intermediate and advanced participants alike, the role play method induces the use of “fillers” such as “oh”, “well”, and “um” more than the DCT method. The difference between the close role play and open role play method lies in that the open role play allows the participants

more turns to express their refusing message. Therefore, the participants tend to come up with different reasons/excuses, which was induced by the interlocutor's feedback, in the second turn of response to strength their refusing intention. Take the refusal response of the intermediate participant #30 in Table 10 as an example. Participant #30 refused the boss's invitation in the 1st turn of response by saying, "I'm sorry, this Sunday? I have to go home to see my mom. This Sunday is my mom's birthday." When the boss tried to invited her again to just stop by for 20 minutes, the participant stated another reason-"But my home is very very long distance. I got to drive three or four hours to get there."- to refuse boss's invitation.

Discussion and Conclusion

The present study compared DCT, closed role play and open role play as data elicitation method of L2 learners' pragmatic competence. The results of the present study showed that the three methods elicit similar samples in terms of response length, mean number of semantic formulas, frequency of semantic formulas and the combination pattern of the semantic formula when comparing the DCT with the first turn of role play, which supports Kasper's observation that "production questionnaires are useful to inform about speakers'pragmalinguistic knowledge of the strategies and linguistic forms by which communicative acts can be implemented and about their sociopragmatic knowledge of the context factors under which particular strategic and linguistic choices are appropriate" (p.329).

The finding of the present study, however, contrasted with the previous study conducted by Beebe and Cumming (1996) which found that DCTs did not elicit natural speech with respect to actual wording, range of formulas and length of responses. One possibility for the inconclusive finding may be because of the research design. Beebe and Cummings (1996) used between-subject design whereas the present study employed the within-subject design. As aforementioned, the use of

different subjects with different tasks may introduce confounding variable in the study. Another possibility may be because of the participants' language proficiency. Beebe and Cummings (1996) collected data from the native speakers of English whereas the present collected data from the non-native speakers of English. As Gass and Houck (1999) pointed out, the speech act of refusal may have greater complexity than other speech acts. It may not be easy for the nonnative speakers with limited linguistic proficiency to refuse in a foreign language with various styles or strategies.

Note that open role play task differs from written DCT and closed role play in that it elicits more turns of refusing responses and allows researchers to examine the sequence of refusal realization strategies. In addition, role play task, closed or open, provided researchers with the information concerning the emotion, pronunciation and intonation when refusing, which cannot be gathered using written DCT. Moreover, the role play task requires the spontaneous oral response from the participants. Whether the spontaneous nature of the role play task provokes more errors than written DCT in their refusal responses warrants further investigation.

Reference

Bardovi-Harlig, K., Hartford, B.A.S., Mahan-Taylor, R., Morgan, M. J. & Reynolds, D.W. (1991). Developing pragmatic awareness: Closing the conversation. *ELT Journal, 45,* 4-15.

Bardovi-Harlig, K., & Hartford, B.A.S (1993). Refining the DCT: Comparing open questionnaires and dialogue completion tasks. In L. Bouton and Y. Kachru (Eds.), *Pragmatic and Language Learning 4*. University of Illinois at Urbana-Champaign.

Beebe, L.M. & Cummings, M. (1985). Speech act performance: A function of the data collection procedure? Paper presented at TESOL '85, New York.

Beebe, L.M. & Cummings, M. (1996). Natural speech act data versus written questionnaire data: How data collection method affects speech act performance. In S. Gass & J. Neu (Eds.), *Speech Acts across Cultures: Challenges to Communication in a Second Language,* pp. 65-86. Berlin: Mouton de Gruyter.

Billmyer, K. & Varghese, M. (2000). Investigating instrument-based pragmatic

Linguistics, 21(4), 517-552.

Cohen, A. (1996). Investigating the production of speech act sets. In S. Gass & J. Neu (Eds.), *Speech Acts across Cultures: Challenges to Communication in a Second Language*, (pp. 21-43). Berlin: Mouton de Gruyter.

Cohen, A. & Olshtain, E. (1994). Researching the production of speech acts. In E. Tarone, S. Gass, and A, Cohen (Eds.), *Research Methodology in Second Language Acquisition*, (pp.143-155). Hillsdale, NJ: Lawrence Erlbaum Associates.

Dahl, M. (1995). *Authentic and role play interaction: A study of method*. Unpublished doctoral dissertation, University of Copenhagen.

Felix-Brasdefer, J.C. (2007). Natural speech vs. elicited data: A comparison of natural and role play requests in Mexican Spanish. *Spanish in Context, 4*(2), 159-185.

Gass, S. and N. Houck. (1999). *Interlanguage Refusals: A cross-cultural study of Japanese-English*. Berlin: Mouton, The Hague.

Golato, A. (2003). Studying compliment responses: A comparison of DCTs and recordings of naturally occurring talk. *Applied Linguistics, 24*(1), 90-121.

Hartford, B.S., & Bardovi-Harlig, K. (1992). Experimental and observational data in the study of interlanguage pragmatics. In L. Bouton and Y. Kachru (Eds.), *Pragmatics and Language Learning*, Monograph Series (3), 33-52. University of Illinois.

Houck, N. & Gass, S. (1996). Non-native refusals: A methodological perspective. In S. Gass & J. Neu (Eds.), *Speech Acts across Cultures: Challenges to Communication in a Second Language*, 46-64. Berlin: Mouton de Gruyter.

Johnston, B., Kasper, G. & Ross, S. (1998). Effect of rejoinders in production questionnaires. *Applied Linguistics, 19*(2), 157-182.

Kasper, G. & Dahl, M. (1991). Research methods in interlanguage pragmatics. *Studies in Second Language Acquisition, 13*(2), 215-247.

Kasper, G. (2000). Data collection in pragmatics research. In H. Spencery-Oatey (Ed.), *Culturally speakin: Managing rapport through talk across cultures*. London and New York: Continuum. Pp.316-41.

Larsen-Freeman, D., & Long, M.H. (1991). *An Introduction to Second Language Acquisition Research*. New York: Longman.

Myers-Scotton, C. & Bernstein, J. (1988). Natural conversation as a code for textbook dialogue. *Applied Linguistics, 9*, 372-384.

Olshtain, E. & Cohen, A. (1991). Teaching speech act behavior to nonnative speakers. In M. Celce-Murcia (Ed.), *Teaching English as a Second or Foreign Language* (2nd edition, pp. 154-190). New York: Newbury House.

Rintell, E.M. & Mitchell, C.J. (1989). Studying requests and apologies: An inquiry into method. In S. Blum-Kulka, J. House, & G. Kasper (Eds.), *Cross-cultural Pragmatics: Requests and Apologies* (pp. 248-272). Norwood. NJ: Ablex.

Robinson, M.A. (1992). Introspective methodology in interlanguage pragmatics research. In G. Kasper (Ed.), *Pragmatics of Japanese as Native and Target Language* (Second Language Teaching and Curriculum Center Technical Report No. 3, pp. 27-82). Honolulu: University of Hawaii Press.

Rose, K. (1992). Speech acts and questionnaires: The effect of hearer response. *Journal of Pragmatics, 17,* 49-62.

Rose, K. (1994). On the validity of discourse completion tests in non-Western contexts. *Applied Linguistics, 15,* 1-14.

Rose, K. & Ono, R. (1995). Eliciting speech act data in Japanese: The effect of questionnaire type. *Language Learning, 45,* 191-223.

Sasaki, M. (1998). Investigating EFL students' production of speech acts: A comparison of production questionnaires and role plays. *Journal of Pragmatics, 30,* 457-484.

Turnbull, W. (2001). An appraisal of pragmatic elicitation techniques for the social psychological study of talk: The case of request refusals. *Pragmatics, 11*(1), 31-61.

Wolfson, N. (1989). *Perspectives: Sociolinguistics and TESOL*. Cambridge, MA:Newbury House.

Yamashita, S. (1996). *Six measures of JSL pragmatics*. (Technical Report #14.) Manoa, HI: University of Hawaii Press.

Yuan, Y. (2001). An inquiry into empirical pragmatics data-gathering methods: Written DCTs, oral DCTs, field notes and natural conversations. *Journal of Pragmatics, 33,* 271-292.

Table 1. Studies comparing different data elicitation methods.

Studies	Between/within subject design	Speech act	Subjects	Data elicitation methods
Beebe & Cummings (1996)	Between	Refusal	NS of English	DCT vs. telephone conversation
Billmyer & Varghese (2000)	Between	Request	NS vs. NNS	Modification of DCT prompt
Dahl (1995)	Between	Disagreement & refusal	NS	Open role play vs. observation
Hartford & Bardovi-Harlig (1992)	Between	Refusal	NS vs. NNS	DCT vs. advising sessions

Hinkel (1997)	Between	Advice	NS vs. NNS	DCT vs. multiple choice questionnaire
Rintell & Mitchell (1989)	Between	Request & apology	NNS	Written vs. oral DCT
Rose (1994)	Between	Request	NS vs. NNS	DCT vs. multiple choice questionnaire
Rose & Ono (1995)	Between	Request	NS of Japanese	DCT vs. multiple choice questionnaire
Yamashita (1996)	Within	Request	NNS	6 measures developed by Hudson et al., 1995
Yuan (2001)	Between	Compliment	NS of Chinese	Written, oral DCTs, field notes and natural conversation

Table 2. Classification of the refusal situations

stimulus type	refuser's social status	social distance	situation
request	equal	unfamiliar	lending notes
invitation	low	unfamiliar	refuse party invitation

Table 3. Means and standard deviations of the response length in DCT, close role play and 1st turn of the open role play for each group

Situation	Proficiency	Task					
		DCT		Close role play		1st turn of open role play	
		M	*SD*	*M*	*SD*	*M*	*SD*
Situation1	*Intermediate*	14.53	5.90	15.93	6.67	12.87	6.27
	advanced	18.30	7.66	18.53	11.41	21.10	11.10
Situation2	*Intermediate*	17.17	7.27	17.50	7.05	15.90	7.96
	advanced	24.63	12.36	26.40	15.40	24.80	12.22

Table 4 Analysis of variance for response length

Source of variance	*df*	*SS*	*MS*	*F*	*p*

Situation 1					
Task	2	21.011	10.506	.146	.864
Proficiency	1	1065.800	1065.800	14.851	.000
Task x Proficiency	2	265.233	132.617	1.848	.161
Situation 2					
Task	2	79.300	39.650	.337	.714
Proficiency	1	3192.022	3192.022	27.152	.000
Task x Proficiency	2	20.544	10.272	.087	.916

Table 5. Means and standard deviations of the response length in the 1st turn and 2nd turn of open role play

Situation	Statistics	Task		*t*-value	*p*
		1st turn of open role play	2nd turn of open role play		
Situation 1	*M*	16.98	15.05	1.49	.141
	SD	9.56	9.86		
Situation 2	*M*	20.35	15.05	4.29	.000
	SD	9.56	11.17		

Table 6. Means and standard deviations of the number of semantic formula used in DCT, close role play and 1st turn of the open role play for each group

Situation	Proficiency	Task					
		DCT		Close role play		1st turn of open role play	
		M	*SD*	*M*	*SD*	*M*	*SD*
Situation1	*Intermediate*	2.67	.80	2.50	.78	2.60	.81
	advanced	2.70	.65	2.77	.86	2.57	.77
Situation2	*Intermediate*	2.87	.82	2.63	.67	2.57	.73
	advanced	2.93	.91	3.37	1.19	3.00	.95

Table 7 Analysis of variance for the number of semantic formula

Source of variance	*df*	*SS*	*MS*	*F*	*p*
Situation 1					
Task	2	.300	.150	.245	.783

Proficiency	1	.356	.356	.581	.447
Task x Proficiency	2	.744	.372	.609	.545
Situation 2					
Task	2	1.411	.706	.886	.414
Proficiency	1	7.606	7.606	9.546	.002
Task x Proficiency	2	3.344	1.672	2.099	.126

Table 8. Means and standard deviations of the pause fillers in DCT, close role play and 1st turn of the open role play

Situation	Proficiency	Task					
		DCT		Close role play		1st turn of open role play	
		M	*SD*	*M*	*SD*	*M*	*SD*
Situation1	*Intermediate*	.17	.38	3.33E-02	.18	3.33E-02	.18
	advanced	.23	.43	.20	.41	6.67E-02	.25
Situation 2	*Intermediate*	.17	.38	3.33E-02	.18	6.67E-02	.25
	advanced	.20	.41	.40	.50	.17	.38

Table 9. Analysis of variance for the number of pause fillers

Source of variance	*df*	*SS*	*MS*	*F*	*p*
Situation 1					
Task	2	.678	.339	3.252	.041
Proficiency	1	.356	.356	3.412	.066
Task x Proficiency	2	.144	7.222E-02	.693	.501
Situation 2					
Task	2	.311	.156	1.168	.313
Proficiency	1	1.250	1.250	9.388	.003
Task x Proficiency	2	.933	.467	3.505	.032

Table 10. Number of combination pattern of semantic formula in DCT, close role play and 1st turn of the open role play

Situation	Task		
	DCT	Close role play	Open role play
Situation 1	29	24	21
Situation 2	26	25	24

Table 11. Most common combination pattern of semantic formula in DCT, close role play and 1st turn of the open role play

	Combination pattern of semantic formula	**DCT**	**Close**	**Open**
Sit. 1	[regret]-[negative ability]-[reason]	9	12	12
	[regret]-[reason]-[alternative]	8	10	6
	[regret]-[reason]*	12	10	14
Sit. 2	[regret]-[negative ability]-[reason]-[alternative]	3	1	6
	[regret]-[negative ability]-[reason]	4	10	5
	[positive opinion]-[reason]	8	5	6
	[regret]-[reason]*	8	9	14
	[regret]-[reason]-[alternative]	3	2	3

Table 12. Examples of refusal responses for DCT, closed role play and open role play for each group.

Situation 2: Refusing boss's invitation to the party:	
President: Hello. My husband and I are very interested in Chinese culture, and have always wanted to talk with you. We are having a little party this Sunday. I wonder if you could come.	
Intermediate participant #30	**Advanced participant # 5**
DCT	**DCT**
"I hope I can come. But that day, I have another date with my boyfriend. I'm sorry"	"I'm so glad that you can invite me, but I have an important meeting tonight, and I am the leader of the meeting. So, I am sorry.
Close role play	**Close role play**
"Oh, sorry. This Sunday is my mom's birthday. I have to come home to celebrate."	"Well, I'm so glad that you invite me to your party. I always want to know you and talk with you. But tonight my church really have an important meeting, and I am the leader of the meeting. I really can't go. Maybe next time, I can find a time, we can go out to have lunch and just talk whatever we want. Will it be okay?"
Open role play	**Open role play**
1st turn response: I'm sorry, this Sunday?	**1st turn response:** It's such a good party.

I have to go home to see my mom. This Sunday is my mom's birthday. **President: Could you stop by for 20 minutes? Just 20 minutes.** **2nd turn response:** But my home is very very long distance. I got to drive three or four hours to get there.	Of course I really want to go. But tonight my church really have an important meeting, and I am the leader. I can't go. I'm so sorry about it. Maybe we can have other time, just go out to have lunch. Talk freely. **President: Could you stop by for 20 minutes? Just 20 minutes.** **2nd turn response:** But after the meeting, there are still other activities. I should stay there to see what's going wrong or things like that. So I really can't go. When the meeting and activities is finished, maybe it's midnight. Sorry.

Appendix A.

Situation 1:

You attend classes regularly and take good notes. Your classmate (an acquaintance) often misses a class and asks you for the lecture notes.

Classmate: Oh God! We have an exam tomorrow but I don't have notes from last week. I am sorry to ask you this, but could you please lend me your notes once again?

You:_____

Situation 2: You work part-time for an American company. The company is very big, and you have never talked to the president (a female) before, but one day she comes by, and invites you to come to a party at her house the following Sunday. She is interested in Chinese culture and wants to talk to you.

President: Hello. My husband and I are very interested in Chinese culture, and have always wanted to talk with you. We are having a little party this Sunday. I wonder if you could come.

You:_____