A Thought on Data Collection Methodologies and Different Kinds of Data in Pragmatics: In Reply to Haugh (2005)

Saeko FUKUSHIMA

1 . Introduction

There are different methodologies for collecting different kinds of data in pragmatics. To my knowledge, Kasper's (2000) review is probably the most extensive of the data collection methodologies most commonly used. She has included spoken interaction, questionnaires, interviews, diaries and think-aloud protocols. Under the category of spoken interaction, authentic discourse, elicited conversation and role-play are included. Production questionnaires, multiple choice and rating scales were included in questionnaires. These different methodologies elicit different kinds of data. Diaries, interviews, and scaled response instruments all elicit *self-report data* (author's emphasis) in isolation from the contexts in which the reported event occurs (Kasper 2000 : 336). Think-aloud protocols (TAP) are *verbalization of thought processes* (author's emphasis) during engagement in a task (Ibid.), which are different in nature from other data such as authentic discourse data or linguistic realization data.

According to Kasper (2000 : 318), authentic data are collected through taking field notes and audio-video-recording. In addition to these approaches, through e-mail messages, which are written, authentic data were collected in Fukushima (2004). In most cases, authentic data have been dealt almost identical with spoken data. However, as in the study by Fukushima (2004), authentic data can be written. Authentic data are, therefore, spoken or written. Elicited linguistic realization data are also sometimes spoken and sometimes written. Spoken linguistic realization data can be elicited through conversation tasks and role plays (Kasper, 2000 : 320-325). Written linguistic realization data can be elicited through questionnaires or multiple choice questionnaires (MCQs). Sometimes spoken data are elicited through the written medium such as production questionnaires or discourse completion tests (DCTs). Although Cohen (2005 : 283) defends the use of DCTs, here arises a limitation of data collection, that is, the spoken data elicited through the written medium lack the naturalness, as the subjects are in most cases asked to write down what they would say under some situations, which may be different from what they would actually say in normal settings.

Multiple choice questionnaires (MCQs) sometimes elicit spoken data through a written medium (e.g., Rose & Ono, 1995). They can also elicit other kinds of data, such as strategy data. Different kinds of data can be collected through different methodologies.

In Fukushima (2000), I have collected strategy data. Such data are more abstract than linguistic realization data and they are different from self-report data, as they are not in isolation from the contexts. In order to elicit strategy data, I have used a multiple choice questionnaire, which has been often used to elicit linguistic realization data in previous studies. As far as I know, strategy data were not elicited through a multiple choice questionnaire in previous research, and strategy data were not the focus of the previous research.

Haugh (2005: 161) in his review of Requests and Culture by Fukushima (2000) pointed out what he considered to be the limitations of the methodology. This paper attempts to clarify the use of the methodology used in Fukushima (2000) and to show the importance of strategy data.
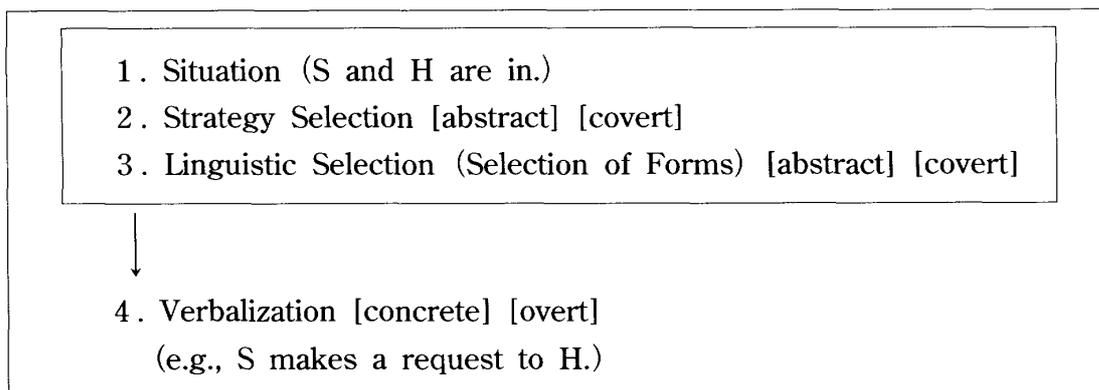
2. Strategy Data

When a speaker (S) makes linguistic acts to a hearer (H), making a request to H, for instance, I believe that there are the following stages (See Figure 1.):

1. S and H are in a certain situation.
2. S selects a strategy.
3. S puts a strategy into words.
4. S verbalizes the linguistic selection.

In Fukushima (2000), I have focused on stage two in the sequence outlined above: strategy selection.

Figure 1. The Process of Linguistic Acts

1. Situation (S and H are in.)
2. Strategy Selection [abstract] [covert]
3. Linguistic Selection (Selection of Forms) [abstract] [covert]

4. Verbalization [concrete] [overt]
   (e.g., S makes a request to H.)

In the first stage, S and H are in a certain situation in which such social variables as power and distance between S and H are present. Such variables as the degree of imposition of the requested act occur when S wants to make a request. Considering such variables, S chooses a strategy in the second stage. S may choose a strategy unconsciously on many occasions, but S may do so consciously when there is a marked item in the situation involved. For example, when the degree of imposition of the requested act is extremely high, S may try to choose a strategy which may compensate for the high degree of imposition by, for example, employing more indirect request strategies than usual. The process of selecting a strategy is covert. In the third stage, S selects appropriate linguistic forms which express the strategy. Then, S verbalizes that, i.e., S says or writes a request to H. Such verbalization is a tangible realization of the covert strategy. Strategy data as such are concerned with the second stage. As Fukada and Asato (2004 : 1997) state, linguistic acts in general occur without the speaker's conscious thoughts. Not only in the second stage, but also in the third stage, S puts a strategy into forms unconsciously in most cases, except for marked occasions. S may not usually be aware that s/he goes through these four stages in order to make linguistic acts.

Some researchers (e.g., Kasper and Dahl 1991; Cohen and Olshtain 1994; Rose and Ono 1995; Sasaki 1998 ; Yuan 2001) have considered methodologies for data collection in pragmatics, but, except for Cohen and Olshtain (1994), most of them have not considered the second strategy selection stage. Cohen and Olshtain (1994 : 146) state that "The process of selecting the socioculturally appropriate strategy and the appropriate sociolinguistic forms for that strategy is complex···." They consider two stages: (1) strategy selection and (2) selection of forms. "Strategy selection" corresponds to the second stage in the above, and "selection of forms" applies to the third stage. Unfortunately, Cohen and Olshtain (1994) do not further discuss the stage of strategy selection. Since the stage of strategy choice cannot be seen, it is difficult to investigate strategy data. This may be why strategy data have not been

the focus of the previous research.

In Fukushima (2000), I have decided to investigate strategy data, despite the difficulty in doing so, because strategy data are at a deeper (or former, i.e., the second stage in Figure 1.) level than linguistic realization data (the fourth stage in Figure 1.). In other words, linguistic realization data are the outcome of strategy data. Therefore, it may be possible to make a more solid comparison between different languages with strategy data than with linguistic realization data. Even if there is some resemblance in linguistic realizations in different languages, there is no guarantee that they mean the same at a deeper level, or they have the same pragmatic meaning. Indeed, Marquez-Reiter (2000) argues that conventional indirectness in English and Spanish does not mean the same with regard to compliance with requests. By focusing on strategy data rather than on linguistic realization data, the limitations of eliciting spoken data through a written medium can be mitigated.

It seems that Haugh (2005) thought that I used a discourse completion test (DCT) and elicited spoken data, and he (2005 : 161) asked whether written data can be expected to represent what the informants would actually say in those situations, pointing out that among other things written data are unnatural because of the focus on a single turn instead of a number of turns as in natural situations.

In fact, Fukushima (2000 : 133-141) contains a review of the advantages and disadvantages of the methodologies used for data collection in previous pragmatic studies. Aware of the limitations of eliciting spoken data via written medium (e.g., DCTs), I tried not to elicit spoken data with a written questionnaire. Instead, strategy data were elicited, for which a multiple choice questionnaire of strategies was employed (e.g., Stating the reason ＋ Making an indirect request), specifically excluding linguistic realizations. Haugh (2005 : 162) states that:

> ⋯ it is perhaps a pity that there was no triangulation of the written questionnaire data with other sources of data (from naturally occurring data or role plays), as this could have alleviated some doubts as to the validity of the data.

While such triangulation is desirable, in fact the study was not focusing on linguistic realizations which would no doubt have benefited from such methodological triangulation.

The subjects in Fukushima (2000) were given a multiple choice questionnaire with strategy choices and asked to choose a strategy. Since the subjects might not

understand the strategy as specified, an example was also given. In the questionnaire instruction (Ibid.: 303), it was emphasized that the respondent was to choose a strategy. Since linguistic realizations were not elicited, it is not a problem even if the examples may differ from what they would actually say. It must be, however, acknowledged that example sentences may have influenced the subjects' strategy choices, because it is not certain whether the subjects' choices were influenced by the strategies or by the examples. This problem could have been solved or at least alleviated to a certain extent if an interview asking the subjects why they had selected certain strategies had followed. It was not, however, feasible to conduct an interview with one hundred and twenty-one subjects in Britain and one hundred and thirty-three subjects in Japan within a limited time.

In order to validate strategy data in future studies, a qualitative study with a smaller number of subjects may be possible, as it is necessary to balance the validity of the data and feasibility when we conduct research. Or it may be also possible to conduct think aloud protocols (TAP), i.e., verbalizations of thought processes during engagement in a task (Kasper 2000: 336), as we may be able to find out what the subjects are thinking or by what the subjects are being influenced when they choose strategies. Or we may be able to combine TAPs and a retrospective interview. Such procedures would go some way to meeting Haugh's desire for triangulation.

3. Conclusion

Clearly, there is no universal method for data collection in pragmatics, and eliciting strategy data through a MCQ with strategy choices is not an exception. However, I hope I could at least show that, despite its limitations, there is a way of eliciting strategy data, not only linguistic realization data. With strategy data, it is possible to compare the data in different languages at a deeper level than with linguistic realization data. It is hoped that more research on strategy data will be conducted so that there will be more solid comparison among different languages. Researchers in pragmatics need to further develop methodologies for data collection and to adjust methodologies according to their research purposes, as the appropriate methodologies vary according to what kinds of data they would like to obtain. The use of strategy elicitation will be an addition to the repertoire.

References

Cohen, Andrew D. and Elite Olshtain. (1994). Researching the Production of Second-

Language Speech Acts. In *Research Methodology in Second-Language Acquisition,* Elaine E. Tarone, Susan M. Gass and Andrew D. Cohen (eds.), 143-156. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Cohen, Andrew D. (2005). Strategies for Learning and Performing L2 Speech Acts. *Intercultural Pragmatics* 2 (3): 275-301.

Fukada, Atsushi and Noriko Asato. (2004). Universal Politeness Theory: Application to the Use of Japanese Honorifics. *Journal of Pragmatics* 36: 1991-2002.

Fukushima, Saeko. (2000). *Requests and Culture: Politeness in British English and Japanese.* Bern: Peter Lang.

Fukushima, Saeko. (2004). Requests in Japanese: A Study through E-mail Messages. Paper presented at International Conference on Language, Politeness and Gender: The Pragmatic Roots at University of Helsinki on the 4th September, 2004.

Haugh, Michael. (2005). A Review of Requests and Culture by Saeko Fukushima. *Journal of Politeness Research* 1 (1): 160-165.

Kasper, Gabriele and Merete Dahl. (1991). Research Methods in Interlanguage Pragmatics. *Studies in Second Language Acquisition* 13: 215-247.

Kasper, Gabriele. (2000). Data Collection in Pragmatics Research. In *Culturally Speaking: Managing Rapport Through Talk Across Cultures,* Helen Spencer-Oatey. (ed.), 316-341. London: Continuum.

Marquez-Reiter, Rosina. (2000). Conventional Indirectness in English and Spanish: Same or different? Paper presented at the Second International Conference in Contrastive Semantics and Pragmatics at Newnham College, Cambridge on the 13th September, 2000.

Rose, Kenneth and Reiko Ono. (1995). Eliciting Speech Act Data in Japanese: The Effect of Questionnaire Type. *Language Learning* 45 (2): 191-223.

Sasaki, Miyuki. (1998). Investigating EFL Students' Production of Speech Acts: A Comparison of Production Questionnaires and Role Plays. *Journal of Pragmatics* 30: 457-484.

Yuan, Yi. (2001). An Inquiry into Empirical Pragmatics Data-gathering Methods:

Written DCTs, Oral DCTs, Field Notes, and Natural Conversations. *Journal of Pragmatics* 33: 271-292.