

A Study on Japanese Students' Communication Strategies in an Oral Test

Soo-im Lee

Abstract

The objective of teaching spoken language is to improve the ability to interact successfully in the target language, however it is difficult to measure the interactive skill accurately. The scoring of oral ability in interview tests tends to be highly subjective, therefore thorough training given to judges and interviewers is essential for increasing the reliability and validity of the measuring instruments. Lee (1996) found numerous causes for judges' subjectivity in their evaluation of an in-house oral interview test, and it was concluded that the subjectivity of the judges had greatly affected their scoring. Oral interview tests usually have subscales, for example, the Foreign Service Institute oral interview (FSI) has five subscales; accent, grammar, vocabulary, fluency and comprehensibility. These subscales are commonly used in a number of oral tests and comprehensibility is one of the most important subscales when measuring testees' interactive performance. This paper investigates the effect of testees' responses which were effectively used as communication strategies on a comprehensibility subscale in an in-house oral interview test. The initial research questions mainly concerned; 1) the types of testee responses which affect scoring on a comprehensibility subscale of an in-house oral interview tests, and 2) the degree to which testee responses could differentiate their comprehensibility. This study also explores the effect of the testees responses on the perceptions of the judges in the interview test. The first phase of the study was designed as a partial replication of Rost and Ross's study (1991) to investigate the relationship between learner use of asking clarifications and

their language proficiency which was measured in a dictation test. In this phase of the study, through discriminant analysis, six types of listener responses were identified and the use of these responses significantly related to the scores of comprehensibility in the in-house oral interview test. The second phase of the study examined the judges' perceptions toward the testee responses and the results of this phase revealed that the six types of responses had positively affected their perceptions and consequently affected the scoring of comprehensibility.

Literature review

Importance of strategies in language learning

One of the fundamental questions in Second Language Acquisition is whether language proficiency in a second language can be enhanced by formal instruction or whether second language proficiency can be achieved unconsciously and spontaneously: being dependent on naturalistic acquisition of the speech of the surrounding community. Krashen's (1981) well-known contrast of learning and acquisition has triggered controversial discussions and it has generated a considerable body of research seeking to answer the above question. It is likely that learning and acquisition are not mutually exclusive but are rather parts of a potentially integrated range of experience. Many researchers suggest that both aspects of acquisition and learning fasten development of communicative competence (Campbell & Wales, 1970; Canale & Swain, 1980; Hymes, 1972; and Omaggio, 1986). Development of communicative competence requires realistic interaction among learners using meaningful, contextualized language. Learners consciously or subconsciously employ various strategies when they participate in authentic communication. The same phenomenon of utilizing strategies is seen in listeners' activities, in other words, listeners try to respond in various ways to keep the interaction with the interviewer going and these responses are defined as communication strategies (CS).

Communication strategies were first defined by Selinker in his account of

the processes for interlanguage(1972). Ellis(1985) differentiates communication strategies from learning strategies, and he characterizes that CSs are problems-oriented. It is assumed that various CSs are employed by L2 listeners to compensate for their weaknesses in comprehensibility. Ellis explains that CS function as follows: Communication strategies are employed by the L2 listeners because he or she lacks or cannot gain access to the linguistic resources required to express an intended meaning. Strategies are especially important for language learning because they are tools for active, self-directed involvement, which is essential for developing improved proficiency and greater self-confidence. Lists of characteristics of good language learners refer to a variety of strategies, such as taking advantage of practice opportunities, willingly and accurately guessing, handling emotional issues in language learning, consciously developing the L2 as a meaning system and a structure system, and monitoring one's own speech (Rubin, 1975; Stern, 1983). Previous research found that effective L2 learners are aware of the strategies they use and why they use them (Abraham & Vann, 1987; O`Malley & Chamot, 1990). Students who are less successful at language learning are likewise able to identify their own strategies; however, they do not know how to choose the appropriate strategies or how to link them together into a useful "strategy chain" (Block, 1986; Galloway & Labarca, 1991; Green & Oxford, 1995; Stern, 1975; Abraham & Vann, 1987).

Passive and active skills

The traditional notions of the four macro-skills (reading, writing, speaking and listening) of language teaching have been reviewed in many studies. Traditional notions in language teaching distinguish the four macro-skills into the following two groups: active/productive skills and passive / receptive skills. Speaking and writing are considered to be active, or productive skills whereas listening and reading are said to be passive, or receptive skills. It might be convenient to represent the language skills in this way when considering language usage, but it is not especially helpful, and indeed might be positively misleading, to represent them in this way when considering language use. The

terms productive and receptive refer to the way language is manifested rather than to the way it is realized in communication.

With reference to language usage, it is perfectly true that speaking is an active or productive act. If we think of speaking in terms of language use, however, the situation is rather different. To begin with, an act of communication through speaking is commonly performed in face-to-face interaction and occurs as part of a dialogue or other form of verbal exchange. What is said, therefore, is dependent on an understanding of what else has been said in the interaction. There is a functional situation for each utterance and contextualization which cannot be ignored as an important element behind the utterance. If for example, something being said in the conversational discourse, the utterance made is not an isolated remark, on the contrary it is based on what has been said previously to the speaker, but will in some way derive from the speaker's understanding or interpretations of what other people have already said. Such listening activity is considered as an active and productive skill in the exact same sense as speaking skill.

Listening skills has been classified into two main types. The first type called interactive listening occurs in conversational discourse, and the listener has opportunities to interact with the speaker. Even though he misses the information which the speaker sent, he would not have to give up but rather can access the same information through the use of various responses. Oral tests can measure this type of listening and the listener can be given chances to clarify unknown meanings, with the result that communication occurs. On the other hand, the second type of listening called non-interactive listening is another listening activity in which the listener is obliged to understand the whole meaning based only on a given context. He is not given any chances to clarify the meaning of the speaker's utterances, and once he misses a piece of information, communication fails.

Understanding spoken language is essentially an inferential process based on a perception of cues rather than a straightforward matching of sound to meaning. Even in non-interactive listening listeners go through the inferential process; that means the listener must find relevant links between what is heard and those aspects of context that might motivate the speaker to make a

particular utterance at a particular time. Meaning is created only by an active listening in which the linguistic form triggers interpretation within the listener's background and in relation to the listeners' purpose, rather than conveying information.

Communication strategies in listening

Although listeners make efforts to make meaning out of the content heard based on the various information through inferential process, there are some situations that listeners fail to make meaning out of the context. Rost (1990) introduces four types of understanding; 1) acceptable understanding, 2) targeted understanding, 3) non-understanding, and 4) misunderstanding. Non-understanding refers to the listener being unable to draw any appropriate inference based on what a speaker has just said. Misunderstanding refers to a conflict between the type of inferences that the speaker had expected the hearer to draw from the speaker's utterances and those inferences that the hearer actually has drawn. Acceptable understanding refers to inferences drawn by a listener that are satisfactory to both speaker and listener. Targeted understanding refers to a specific interpretation received that was intended by the speaker.

In a study of listener construct of meaning Rost and Ross (1991) classify L2 learner use of listener feedback, particularly their use of clarification questions in NS-NNS discourse into the four move types. These move types are defined as global reprise, local reprise, forward inferencing and continuation signals. Eight strategies observed in the interaction between narrators and their Japanese interlocutors were entered into a linear discriminant analysis. In the discriminant analysis, the univariate F tests indicated that forward inferencing and continuation signals are strategies that "high" proficiency listeners use and lexical reprise and global reprise are more likely to be used by "low" proficiency listeners. Their findings clearly show that the degree of L2 learners' understanding or construction of meaning are reflected by the frequency of listener feedback strategies.

Other studies on listening skills and strategies

Fujiwara (1990), in a study of reporting on Japanese EFL students' listening skills and strategies concluded that:

In general, the good learners use a wide variety of materials and strategies. They are able to analyze what they are learning, how they are learning it, and what they need to learn. They can assess their own areas of strengths and weaknesses and the progress they have made. They find ways to overcome the difficulties they encounter. They show creativity in designing tasks and diligence in carrying them out (1990:23-24).

Varonis and Gass (1985) investigated conversational interactions between native speakers (NS)/ non-native speakers (NNS) and NNS/NNS conversations, noting that negotiation of meaning is most prevalent among NNS/NNS pairs. The types of negotiation of meaning in their study are regarded as CSs. They suggest that communication between any two interlocutors is facilitated when they share a common background, linguistic or otherwise. Even so, as a number of investigators note (Grosz 1979; Hobbs & Robinson 1978; Labov & Fanshel 1977; Marlos 1981), interlocutors also need to share an understanding of each other's goals in the interaction for the conversation to proceed with maximum ease. Among non-native speakers, however, even when interlocutors realize each other's goals, other factors, for example linguistic, social, or cultural factors, may lead to a breakdown in communication. Such breakdowns must be negotiated before a further exchange of information can proceed. In addition, negotiation may strengthen the social or interpersonal dimension of an interaction by acting as a conversational continuant.

Paribakhat's study (1985) of listener and speaker negotiation focused on strategic competence and examined the relationship between speakers' proficiency level in the target language and their communication strategies use. The results of the study show that in the majority of the comparative analyses, both in terms of types and relative frequency of use of CSs, the advanced learners were in a mid-position between the native speakers and the

low-proficiency learners. The data suggest that learners' use of CSs had specific characteristics at different developmental stages of their interlanguages. That is, learners seem to abandon or adopt certain CSs, and also altered their proportional use of certain strategies as they approached the target language. Learner behaviors in terms of strategy use seems, therefore, to be transitional and dynamic. In that all subjects groups adopt basically the same types of CSs, it would appear that these speakers share a certain ability or competency, referred to as strategic competence. This is consistent with the position put forward by Canale and Swain (1980), in which they refer to strategic competence as one of the component competences of communicative competence.

The purpose of the study

As cited above, studies of face-to-face interaction have tended to concentrate more on the speaker's role than on the listener's role. Few researchers, however, have made extensive inquiries into how listeners respond in conversational interaction, and yet it is believed that a listener response plays an important role in successful interactions. This is the reason why a comprehensibility subscale should be included in all oral interview tests. If an oral interview test is designed to measure the testees' interactive ability, active listening skill is expected from testees as well as a high level oral proficiency. The purposes of this study are as follows;

1. to identify what types of responses were utilized as CSs by testees in their discourse with interviewers,
2. to investigate the relationship between the testees' CSs and comprehensibility in an in-house oral interview tests,
3. to investigate how the judges perceived the testees' CSs and the degree to which their perceptions affected the result of their scoring in comprehensibility.

The aforementioned research questions were restated in the following hypotheses:

- H1: The types of testees' responses will significantly affect the scores of comprehensibility in an in-house oral interview test.
- H2: To which degree would the choice of testees' responses used as CSs discriminate the three levels of their comprehensibility (High, Intermediate, and Low) in the in-house oral interview test?
- H3: The choice of CSs will significantly predict the scores of the comprehensibility in OIT.
- H4: The choice of CSs will significantly effect the perceptions of the judges in OIT.

Hypothesis 1 was tested with multivariate test and univariate test, and Hypothesis 2 was examined through discriminant analysis. Hypothesis 3 was investigated with multiple linear regression. The three hypotheses are discussed in the phase 1 of this study, and the hypothesis 4 was tested qualitatively adding the descriptive statistics, and it is discussed in the phase 2 of the study.

Methodology

Subjects

Two hundred and seventy Japanese students studying English conversation at a Japanese language school participated in this study as subjects. They currently study English conversation at this private language school and the subject group consisted of 246 university students, 14 high school students and 10 non-students. The students who entered this institute from March 1995 to September 1996 were chosen as the subjects in this study, therefore the subject group in the present study is an intact group. The subjects were divided into three groups of High Listening Proficiency (HLP), Intermediate Listening Proficiency (ILP), and Low Listening Proficiency (LLP) according to the scores of comprehensibility subscale in an in-house oral interview test.

Phase 1: primary data

An oral interview test (OIT) used as one section of the placement test at

the private language school was used to measure the subjects' interactive listening skill. The test was originally designed by the institute and it was used to diagnose the students' strengths and weaknesses in their oral performance. Testees are interviewed by native speakers and each test is 15 minutes long. The interviewer takes the role of the judge, as well. The purpose of the OIT is to measure how successfully the students can continue to interact with a native speakers in English. The questions in the interview are made based on their personal information which is filled out in the questionnaire prior to the interview. The rating of the tests was normed by obtaining a mutual agreement among the seven judges about the values in the scale and evaluating a sample taped interview beforehand to achieve higher interrater and intrarater reliability. The scores were determined according to five categories: pronunciation, fluency, vocabulary, grammar and comprehensibility. Each component has a 5 point scale (1 for very poor, 2 poor, 3 for adequate, 4 for good and 5 for excellent) and the maximum score is 25. All interview tests were video-taped and the tapes were observed by the testees and the school staff, and after the test they discussed the testees' strengths and weaknesses for diagnostic purposes. The subjects were divided into three groups according to the scores of the comprehensibility subscale in the OIT (the the subjects who scored 4 and 5 for HLP, 3 for ILP, and 1 and 2 for LLP.)

The reliability of the oral test

Thirty students' oral performance was video-taped to measure the reliability of the measuring instruments. The tape was shown to two different judges and the two sets of scores were statistically analyzed by the Pearson product moment correlation. The correlation coefficient was .75 and the standard error of measurement was 0.7. This can be interpreted directly as the proportion of the score variance that is consistent. In other words, approximately 75% of the variance among scores can be considered consistent variance, while the remaining 25% should be reviewed as random error variance. A strategy for training the other five raters was developed by the first two judges and a scoring guide with explanation of the assessment process and ex-

ample tapes of OIT were used for practice in assigning ratings.

The initial part of the tape was omitted from this study because the students acted nervously in the beginning of the OIT. From the middle to the end part of each interview, a total of 5 minutes was extracted for this study. The researcher employed five administrative staff whose language proficiency was advanced (above 550 TOEFL scores (Test of English as a Foreign Language) to assist her for this study, and they were instructed to categorize and count all instances of listener responses identified during the 5 minutes interaction between the testees and interviewers in the OIT. These five assistants were thoroughly trained to count and categorize listener responses in order to achieve a high reliability of the measuring instrument. The interrater reliability of this measuring instrument was .82.

The taxonomy

Various listener responses were identified and these were categorized quite differently from the categorization in Rost and Ross's study (1991). The relationship between listener clarification questions and proficiency level was investigated and analyzed in their study, on the other hand, comprehensibility in the oral interview test was a main focus in this study. Therefore, the following variables were originally operationalized and also the categorization was made under each of the understanding stages defined by Rost (1990).

1. Non-understanding stage

a. Maintenance strategy

The listener sends a message that he/she doesn't understand the question or statement so that he/she wants the speaker to repeat the same utterance by maintaining the conversation rather than moving neither backward or forward. Some responses such as "Eh?" or "Hun?" were made in Japanese.

NS: So, how do you like your school?

NNS: Eh?

Expressions of "Pardon me?" "Would you repeat it again?" "Would you speak more slowly?" were used to maintain the conversation.

NS: I like sushi very much, but I don't like horse radish in it.

NNS: **Pardon** me?

2. Incomplete understanding

b. Insisting strategy

The listener sends a message that his/her understanding is not enough. Therefore, he/she asks for clarifications of the question or statement by insisting on clarification of the statement rather than impelling the conversation move forward.

NS: I decided to stay in the same apartment, but the landlady wanted to raise the rent. I couldn't afford to pay 15man (15,000) Yen all by myself. So, I moved to a new one in Mino.

NNS: **What? What about the apartment?**

NS: I send my kids to an international school in Kyoto. The tuition there is outrageously expensive because the Ministry of Education doesn't give any financial support to the school. Why? Because they don't follow the Japanese curriculum, you know, they've got an American system.

NNS: **Who? Who don't** help the international school?

c. Repeating strategy

Repeating or rephrasing specific parts of an utterance with rising intonation to indicate the listener didn't understand and is asking for clarification.

NS: My father was a social worker.

NNS: **social?**

NS: Yes. social worker. Do you know what social workers do?

3. Understanding stage

d. Continuation strategy

The listener understands the utterance of the speaker so that he/she re-

sponds with expressions which indicate that the listener wants the speaker to continue the conversation without taking a turn. Various backchannelling devices were used as this type of strategy.

NS: I finally got used to eating raw fish and now I can eat almost everything except octopus.

NNS: **HU::M**

NS: I came to Japan 7 years ago as an exchange student at Kansai Daigaku.

NNS: **Oh**, I see.

NS: I love Japanese food. Many foreigners can't eat raw fish, but I love them.

NNS: Oh, that's **wonderful**.

NS: I love Japanese movies and my favorite is TORA san.

NNS: **Really?**

NS: Ya, Do you like TORA san, too?

NNS: Not so much.

e. Inferencing strategy

The listener uses this strategy to confirm his/her understanding by paraphrasing the question or utterance.

NS: I cut down my weight almost 4 kilo grams after stopping eating junk food.

NNS: You mean you stopped eating bad food for **health**.

NS: That's right. For example, I stopped eating potato chips and sweets. I used to eat them every day before.

NS: I've just got CATV. It's great. I can watch CNN for 24 hours, and they show many movies.

NNS: **What TV?** Are you **talking** about cable television?

f. Allowing turn taking strategy

After the listener answers the question, he/she initiates the conversation by taking the next turn and introduces a new topic.

NS: You like reading books. What kinds of books do you like to read?

NNS: I like detective stories, especially Agasa Cristy's books are so interesting.

What about your hobby? Do you like sports?

NS: Oh, yes. I play baseball with my friends every month.

These six types of listeners' responses were identified and the frequency of these instances was counted for descriptive statistics.

Results

Table 1 shows the six categorized listener responses which appeared in the OIT. The frequency data was collected and analyzed. Coefficients of skewness and kurtosis of all six variables fell between -1 and 1 so that the distributions of these independent variables were assumed to have normal distributions.

TABLE1

Descriptive statistics of six types of responses

Types	N	Min	Max	M	SD	Skewness	Kurtosis
MA	270	0	5	1.5	1.2	0.30	-0.93
IN	270	0	7	2.7	1.7	0.42	-0.15
RE	270	0	6	1.9	1.4	0.44	0.07
CO	270	0	8	2.8	2.3	0.48	-0.96
INF	270	0	6	1.5	1.5	0.62	-0.80
TU	270	0	5	1.9	1.2	0.13	-0.65

MA=Maintenance strategy, IN=Insisting strategy, RE=Repeating strategy

CO=Continuation strategy, INF=Inferencing strategy, TU=Allowing turn taking strategy

Table 2 shows the descriptive statistics of the three groups (HLP, ILP, LLP) on the instances identified in the OIT. The means of the Maintenance strategy showed the minimum range between the three groups (1.7 /LLP, 1.6/ILP, 1.1/HLP). The means of the Continuation strategy showed the largest range between the three groups (1.2/LLP, 2.2/ILP, 4.1/HLP).

TABLE 2

Means of the six types in the three groups of comprehensibility

Level	MA	IN	RE	CO	IN	TU
LLP	1.7	1.4	1.4	1.2	0.6	1.3
ILP	1.6	2.2	1.6	2.2	1.2	1.6
HLP	1.1	3.6	2.4	4.1	2.2	2.4

LLP=Low listening proficiency, ILP=Intermediate listening proficiency, HLP=High listening proficiency, MA=Maintenance strategy, IN=Insisting strategy, RE=Repeating strategy
CO=Continuation strategy, INF=Inferring strategy, TU=Allowing turn taking strategy

Six types of responses observed in the OIT between testees and interviewers were analyzed through univariate tests. The main effect of the dependent variables; the six types of testee responses on the independent variable; scores of comprehensibility. The multivariate results are displayed in Table 3 and the test results are highly significant at the $p < .000$ level. The result of the subsequent univariate F tests for each dependent variable also show significance and confirms that the use of CSs significantly discriminate the group memberships.

TABLE 3(1)

Multivariate F tests

	Value	P-level
Wilks' Lambda	0.6503	
Rao R	10.4836	.000*
Pillai-Bartlett Trace	0.3532	
V	9.4012	.000*

* $p < .01$

TABLE 3(2)

The Results of Univariate F tests

	SS	Error	df	F	P
MA	9.00	1.49	2,267	6.05	0.002*
IN	82.30	2.45	2,267	33.55	0.000*
RE	25.95	1.67	2,267	15.48	0.000*
CO	149.48	4.12	2,267	36.20	0.000*
INF	47.99	2.01	2,267	23.80	0.000*
TU	26.49	1.18	2,267	22.26	0.000*

* $p < .01$

MA=Maintenance strategy, IN=Insisting strategy, RE=Repeating strategy
CO=Continuation strategy, INF=Inferring strategy, TU=Allowing turn taking strategy

Discriminant function analysis is used to determine which variables discriminate between two or more naturally occurring or pre-selected groups. In this study, three levels (HLP, ILP, and LLP) in the scores of comprehensibility of the OIT were pre-selected for a multiple discriminant function analysis. The purpose of the analysis is to determine a set of characteristics that allows the best discrimination between the three groups. In testing the discriminant analysis, to see which ones contribute to the discrimination between groups. Table 4 shows the Eigenvalues for the two discriminant functions (canonical root) and the percents of explained variance accounted for. These percentages pertain to the variability in the variables that discriminates between groups (The first root accounts for almost 98% of the common variance). Common variance refers to the variability of the dependent variables that can be explained to be the between-group factor. As seen in table 4, the sign of the discriminant function coefficient for variable of Insisting, Repeating and Continuation have relatively heavier weights than the other three variables; Maintenance, Inferencing, and Allowing turn taking strategies. All dependent variables show positive signs except the Maintenance strategy which has a negative sign. The second root accounts for only 2% of the common variance, therefore the magnitude of the discrimination is much smaller.

Table 5 displays the means for the discriminant functions. As it is seen that the first function seems to discriminate mostly between HLP, versus ILP and LLP combined. The second function doesn't seem to discriminate these three groups as distinct as the first function. Since the first root accounts for over 98% of the explained variance, that is, 98% of all discriminatory power is explained by the first function. Thus, the first function is clearly the most important one to be interpreted. The function 1 discriminates mostly between HLP and LLP, however, there is not clear distinctive curve between the LLP and ILP. The function 2 seems to distinguish mostly between ILP and the other two levels, however, the magnitude of the discrimination is much smaller.

TABLE 4

Eigenvalues and percent explained

Variables	Root1(98%)	Root2(2%)
MA	-0.16	0.72
IN	0.45	0.36
RE	0.61	-0.56
CO	0.52	0.23
INF	0.16	0.58
TU	0.12	-0.51

MA=Maintanance strategy, IN=Insisting strategy, RE=Repeating strategy

CO=Continuation strategy, INF=Inferencing strategy, TU=Allowing tum taking stragety

TABLE 5

Means of discriminant functions

Groups	discriminant function 1	discriminant function 2
LLP	1.305	.524
ILP	1.847	.820
HLP	3.19	.705

LLP=Low listening proficiency, ILP=Intermediate listening proficiency,

HLP=High listening proficiency

In table 6 the result of classification matrix show that each group showed inconsistency in choices of strategies, particularly all subjects in LLP were misclassified mostly as ILP group (35 subjects) and 2 subjects were classified as HLP group and the correct percent of cases was 0%. ILP group achieved the highest correct classifications (80.4%) in choice of strategies. 32 HLP subjects were misclassified as ILP, therefore, HLP percent correct was relatively low (66.3%).

TABLE 6

Observed and predicted groups memberships

group	percent correct	L p=.13	I p=.51	H p=.35
LLP	0	0	35	2
ILP	80.4	0	111	27
HLP	66.3	0	32	63
TOTAL	64.4	0	178	92

In the next study multiple linear regression was used to find out which of multiple CSs best predict the raters' scores of comprehensibility subscale. Table 7 displays the correlations between the variables (R and R^2), the standardized regression coefficients (β), the standard errors of β , and the significant F ratios. Continuation strategy proved to be a predictor of achievement for the comprehensibility measure. A stepwise regression indicated that after Step 1, with Continuation strategy in the equation, $R^2 = .27$, $F(1, 268) = 98.94$, $p < .001$. After Step 2, with Insisting strategy in the equation, prediction of the comprehensibility measure improved, $R^2 = .31$, $F(2, 267) = 17.1$, $p < 0.01$. Multiple R is the correlation of each strategy with the comprehensibility scores. The R^2 column shows the amount of overlap accounted for by Continuation strategy; .27. That is, the two measures share 27% variance. It is the best predictor to use in estimating the comprehensibility scores. The next value which is Insisting strategy in the Multiple R and R^2 is .562 or .31. This says Continuation and Insisting strategies together have a 31% overlap with the comprehensibility scores. It shows only a 4% gain in prediction. When Maintenance strategy and Inferencing strategy are added to the equation, the four variables of Continuation, Insisting, Maintenance and Inferencing strategies now account for 34% of the variance in the comprehensibility scores. Another 3% from the last two variables of Maintenance and Inferencing strategies has been added in terms of prediction of scores.

The results of multiple regression share significant finding on the effect of

Continuation strategy with the consequences which discriminant function analysis displayed. Although both Repeating strategy and Insisting strategy which displayed relatively heavier weights in Eigenvalues in the discriminant function analysis, only Insisting strategy showed statistically significant result, but not Repeating strategy in the multiple regression analysis.

Maintenance strategy, Inferencing strategy and Allowing turn taking strategy did not improve prediction of the measure. The Beta coefficients are the values which would be obtained when all of the variables were standardized to a mean of 0 and a standard deviation of 1. Thus, the magnitude of these Beta coefficients display the relative contribution of each independent variable in the prediction of the dependent variable. Continuation strategy shows the largest Beta coefficient (.29, $p < .001$) and Insisting strategy shows the second largest Beta coefficient (.17, $p < .05$).

TABLE 7

Summary of stepwise regression

Variables	Multiple R	Multiple R squar	Multiple R-squar change	F	Beta
CO	0.52	0.27	0.270	98.94*	0.29
IN	0.56	0.31	0.040	17.1*	0.17
MA	0.57	0.33	0.016	6.3*	0.17
INF	0.59	0.34	0.015	6	0.19
TU	0.59	0.35	0.006	2.3	0.10

* $p < .01$

Table 8 shows the correlation matrix among the variables. In the present study, multicollinearity, the term applied to strong interrelationships among predictor variables, was not evidenced in the Pearson product-moment correlation coefficients among all the variables. Therefore, the results of the multiple regression on the CSs are considered legitimate to predict the accomplishments in the comprehensibility measures.

TABLE 8

Correlations

	COM	MA	IN	RE	CO	INF	TU
COM	1	-0.13	0.5	0.11	0.52	0.51	0.43
MA	-0.13	1	-0.35	0.22	-0.45	-0.44	-16
IN	0.50	-0.35	1	0.07	0.65	0.65	0.58
RE	0.11	0.22	0.07	1	0.04	0.04	0.09
CO	0.52	-0.45	0.65	0.04	1	0.72	0.6
INF	0.51	-0.44	0.75	0.03	0.72	1	0.62
TU	0.43	0.17	0.58	0.09	0.6	0.62	1

COM=Comprehensibility score

MA=Maintenance strategy, IN=Insisting strategy, RE=Repeating strategy

CO=Continuation strategy, INF=inferencing strategy, TU=Allowing turn taking strategy

Phase 2 of the study

Twenty students whose listening skill was considered similar were chosen among the first 270 subjects for the second phase of the study. The first group (group A) of ten students consists of those who have been overseas from the period of eight months to 1.2 months. Their purpose of studying abroad was mainly to study English as a second language and all of the subjects went to English speaking countries to study English as a second language (ESL). The second group (group B) of the remaining ten students consist of those who have never been overseas before. Their main method of achieving the language proficiency was based on self-learning or going to private language schools in Japan.

The twenty students' listening skill was measured using Longman, Preparation Course for TOEFL (Test of English as a Foreign Language) TEST, Practice Tests (1996). They took the test of listening section of practice test 1 of the book and their raw scores were converted to converted scores. The mean scores of group A and B were 55 and 54 in converted scores respectively. The two

groups of TOEFL scores were statistically compared and a t-test showed there was no significant difference between the two groups in the listening scores. Their performance for the OIT as the primary data used for the first phase of the study was analyzed to investigate their oral performance.

Results of Phase 2

Table 9 shows the number and types of responses observed in their oral production. The judges who administered the OIT to those 20 subjects were asked to write their impressions on the subjects' performance in the comment section. The comments were carefully analyzed, and the number of words which commonly appeared in the judges' comment section were counted.

TABLE 9

The mean scores and standard deviations of responses

	COM	MA	IN	RE	CO	INF	TU
Group A	4.4(0.7)	0.2(0.4)	1.8(0.9)	0.6(0.8)	5.4(1.0)	3.1(0.8)	2.7(0.3)
Group B	3.8(0.6)	2.3(0.7)	1.2(0.9)	2.4(1.2)	3.0(1.2)	1.3(1.1)	2.0(0.2)

Group A who have some experience of staying overseas achieved a higher group average score in the comprehensibility score than did group B (4.4 and 3.8 respectively.) Group A showed higher frequency data in the Insisting strategy, Continuation strategy and Inferencing strategy and Turn taking strategy while group B showed high frequency in the other remaining variables of Maintenance strategy and Repeating strategy.

All the raters were also asked to write on what impression they received of each of the twenty students within 10 words. The words the judges used for group A were "positive", "friendly", "cooperative", "interactive", "good", "cooperative" and "unlike Japanese" while group B received the comments such as "quiet", "shy", "modest", "nervous", "rigid" and "typical Japanese". Group A seemed to have received more positive comments than did group B in terms

of the judges' perceptions.

Discussion

The present study investigates of the effect of testee's responses in the OIT and their relationship to the scores on a comprehensibility subscale of the test. Before the results of this study are analyzed for discussion, it is important to consider the nature of oral interview tests. Oral tests are usually conducted in a face-to-face interview test situation, the subjectivity of the interviewer is reflected in both areas of procedure and evaluation. The scoring, particularly in interview tests, tends to be more impressionistic than scoring in other types of tests. Compared comprehensibility in interactive tests such as oral interview tests with listening ability in non-interactive tests, the validity of these two tests are apparently different. Comprehensibility in oral tests are measured based on testees' responses to what they were previously told by an interviewer. Judges in oral interview tests decide testees' comprehensibility scores subjectively or impressionistically and the results of this study imply that the testees' responses are the most significant variable which affects their measurement. Non-interactive tests, on the other hand, aim at measuring testees' understanding of what they heard and testees are not given any opportunities to ask for clarifications of the statements. However, testees in interactive tests are given opportunities to exercise their communication strategies to compensate for their weaknesses resulting in a lack of comprehension. It means that comprehensibility scores are closely related to testees' language proficiency, and particularly the discourse strategy (Canale, 1983) which compensates for the listeners' weaknesses greatly effect the comprehensibility score. Therefore, it is interpreted that the validity of comprehensibility in oral interview tests is to measure testees' ability to understand plus an additional skill: interactive ability.

In phase 1 of this study, the results of discriminate function analysis on 270 subjects indicated that the use of certain strategies is significantly correlated with the scores of comprehensibility in an oral interview test. The visual result of discriminant functions clearly shows that LLP has the lowest value

on the function 1, while the HLP has the highest mean; ILP fall in between, but much closer to LLP. In Ross and Rost's study, the two extreme levels; High language proficiency and Low language proficiency were compared in the use of listeners' clarification strategies, and they confirmed that the two proficiency ranges were non overlapping and furthermore there was a clear break between the highest scoring low proficiency listener and the lowest scoring high proficiency listener. On the other hand, In the present study there was no clear break between LLP and ILP, therefore only focusing on the results of LLP and ILP together as one group and HLP are seemingly plausible for analysis.

The results of the univariate F tests and multivariate F tests showed that all the dependent variables significantly discriminated the two groups (LLP and ILP as one group and HLP). The weights in Eigenvalues of the Repeating strategy, Continuation strategy and Insisting strategy were .61, .52, and .45 respectively suggesting that HLP group were positive and collaborative in continuing their conversation with the interviewers using these three responses as communication strategies. In Ross and Rost's study, testees were allowed to ask for clarifications in Japanese in their discourse with the native speaker, and it was found that global reprise (listeners do not understand what the interlocutor had said at all and they ask for clarifications by saying "I don't understand." and lexical reprise (listeners ask the meaning of unknown words) were more likely to be used by low proficiency listeners. In this study, on the other hand, even just repeating an unknown word in order to ask for clarifications was considered more useful than saying nothing to the interviewers. The Insisting strategy, which was operationalized similarly to the low proficiency listener strategy: lexical reprise in Ross and Rost's study, was effectively used by HLP group twice more than did LLP and ILP as one group (the mean for HLP and LLP and ILP as one group are 3.6 and 1.8.) The Continuation strategy was identified as an effective communication strategy employed by HLP group and this finding matches with the Ross and Rost's study. Some subjects in HLP group sent backchannelling signals in Japanese, for example, "Hee!" "Hontoni!" and "Hu::N", however these instances had the same interactive function as the English backchannelling devices and they were effectively

used to continue the conversation flow. The Maintenance strategy showed a negative Eigenvalue and percent suggesting that LLP and ILP as one group show a low discriminant function for this strategy. However, the weight of the Eigenvalue is -.16 so that comparing this variable with the other three variables of the Repeating, Continuation, Insisting strategy, and the Maintenance strategy shows less variability between the groups. The two remaining variables of the Inferencing strategy and Allowing turn taking strategy also indicated small weights in the eigenvalue (.16 and .12 respectively.) Therefore, it is concluded that the Insisting strategy, Repeating and Continuation strategies showed main effects in discriminating the groups.

The results of multiple regression showed that Continuation strategy was the best predictor to the achievement in the comprehensibility scores. It appears that the frequency of Continuation strategy is highly correlated with the comprehensibility scores. Multiple R is the correlation of each strategy with the comprehensibility scores. The results of multiple regression indicate the similar indication made from the results of discriminant function analysis. Good listeners can be defined as cooperative listeners who frequently send effective responses to the speakers; the interviewers in OIT. The cooperative attitude is reflected in the image of the listeners; test takers of OIT and eventually the good image is reflected in the comprehensibility scores.

Phase 2 of this study also found significant effects of CSs regarding the perceptions of the judges. Group A demonstrated higher use of Insisting strategy, Continuation strategy, Inferencing strategy and Turn taking strategy than did the group B. Even though the means of the listening skill for both groups are at similar listening proficiency levels, the two groups performed quite differently in the OIT. Closer observation of their performance revealed that the subjects in group A seemed to enjoy the conversation with the interviewer as if they were not in a test situation while the students in group B kept their rigid and slightly nervous attitude. The students in group A had an experience of living in America from five months to 1.2 years and it is assumed that they have learned how to interact with people in English naturally. However, The students in group B who have never been overseas before learned English only in the classroom situation. The seven judges described their impression

on these two groups in a quite similar way. The content analysis indicated that their perception of group A was positive, friendly, good , interactive, cooperative, good and one of the judges described their performance as "unlike Japanese." On the other hand, the students in group B received rather negative impression from them such as quiet, shy, average, nervous, rigid, and "typical Japanese."

Through language people engage in social interaction with each other. This interaction does not proceed as a rigid series of one-way messages but is the product of collaboration and negotiation. Even in test situations, collaboration and negotiation between testees and judges should be achieved if the validity of the test is to measure how successfully testees can interact with interviewers. Within their shared world, which changes constantly as the interaction proceeds, testees at high level proficiency should be able to interpret, question, clarify and respond to interviewer's messages. Ross and Rost (1991) claim that good learners initiate repair and can stay involved as active participants in interaction, and the same findings were drawn in this study.

Limitations of the study

The reliability of the oral interview test was .75 and the reliability was measured by the classical way of test-retest method using two judges. This reliability result was, however, obtained using two experienced judges. These two judges were also experienced teachers in EFL in Japan; one had three year teaching experience and the other had three and half years experience. Four judges among the seven had less than one year teaching experience, therefore if these novice teachers were tested to obtain the reliability, the result would have been different and possibly would have a lower reliability than .75. In addition to interrater reliability, intrarater reliability should have been measured.

Another factor which might have affected the results was the testees' motivational and attitudinal factors. In phase 2 of this study group A received much positive impression from the judges than did group B, but group A consisted of eight female and two male university students group B consisted of

three female and seven male university students. Gender might have affected the perceptions of the judges and subsequently good attitudinal impression might have led to higher scores of comprehensibility. This motivational and attitudinal factors are worth being investigated for future studies on communication strategies in oral interactive testing.

As mentioned in the discussion section, the discriminant functions showed that LLP and ILP had the similar values and subsequently the two groups were treated as one group comparing HLP group to draw the results. Since the level of comprehensibility was divided into three groups based on the scores of the comprehensibility subscale in the OIT, the scores between LLP and ILP might not be clearly separated (the scores of 1 and 2 for LLP and 3 for ILP). The mistreatment of the raw data might have set strains on the findings of the study. It might have provided more valid results if the two extreme groups of LLP and HLP were extracted for this study and a future study should be conducted following the procedure. Another prospective study will be a study that investigates what variables discriminate testees' comprehensibility focusing on the same language proficiency group.

Conclusion

'Silence' is interpreted in Japanese culture quite differently compared to the norm in western culture. There are many Japanese proverbs about 'silence', for example, "Silence is golden," and "Actions speak louder than words", and these proverbs imply that 'silence' is not considered as a bad manner. Oriental languages represented by Japanese language are defined as a high contextualized language. It means non-verbal factors weight heavily in the communication process and careful listening which concentrates on speakers' talk but other elements in the situations is considered important in Japanese culture. No one way of communicating is better than the other, therefore conforming to Western norms without adequate crosscultural understanding might cause cultural friction to Japanese novice learners. Ross and Rost (1991) concluded in their study that strategies for listening comprehension can be demonstrated and readily adopted. Certain expressions such as asking for clar-

ifications such as "I beg your pardon?", "Would you repeat the question again?" or "Would you speak more slowly?" help learners initiate repairs in understanding and eventually improve their comprehensibility. The findings of this study suggest that teaching these expressions as useful communication strategies in the classroom should be encouraged in order to facilitate students' active listening skill.

References

- Abraham, R. G., & Vann, R. J.(1987). Strategies of two language learners: A case study. In A. Wenden, & J. Rubin (Eds.), *Learner Strategies in Language Learning*. Englewood Cliffs, NJ: Prentice Hall.
- Block, E. (1986). The comprehension strategies of second language readers. *TESOL Quarterly*, 20, 463-494.
- Campbell, R., & Wales, R. (1970). The study of language acquisition. In J. Lyons (Ed.), *New Horizons in Linguistics*. Harmondsworth: Penguin.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics* 1: 1-47
- Ellis, R. (1985). *Understanding Second Language Acquisition*. Oxford: Oxford University Press.
- Fujiwara, B.(1990). Learner training in listening strategies. *JALT Journal*, 12 (2), 203-217.
- Galloway, V., & Labarca, A.(1991). From student to learner: Style, process and strategy. In D. W. Birckbichler (Ed.), *New Perspectives and New Directions in Foreign Language Education*. Lincolnwood, IL: National Textbook Co./ American Council on the Teaching of Foreign Languages.
- Green, M. J., & Oxford, R.(1995). A closer look at learning strategies L2 proficiency, and gender. *TESOL Quarterly*, 29 (2).
- Grosz, B.(1979). Utterance and objective: issues in natural language communication. *Technical Note*, 185, Artificial Intelligence Center; SRI International, California.
- Hobbs, J., & Robinson, J. (1978). Why ask? *Technical Note*, 169, Artificial Intelligence Center; SRI International, California.
- Hymes, D.(1972). On communicative competence. In J.B. Pride & J. Holmes. (Eds.), *Sociolinguistics* (269-293): Selected Readings. Harmondsworth: Penguin.
- Krashen, S.(1981). *Second Language Acquisition and Second Language Learning*. Oxford: Pergamon.
- Labov, W., & Fanshel, D. (1977). *Therapeutic Discourse*. New York: Academic Press.
- Lee, S. (1996). A study of Interrater reliability in an in-house oral proficiency test.

Osakajogakuin Kiyo, 26.

- Marlos, E. (1981). Why answer? A goal-based analysis of a speech event, in D. Sankoff & H. Cedergren (Eds.), *Variation Omnibus*. Carbondale, Ill: Linguistic Research Inc.
- Omaggio, A.C. (1986). *Teaching Language in Context: Proficiency-Oriented Instruction*. Boston, Mass.: Heinle and Heinle.
- O'Malley, J.M., & Chamot, A.U. (1990). *Learning Strategies in Second Language Acquisition*. Cambridge: Cambridge University Press.
- Paribakht, T. (1985). Strategic competence and language proficiency. *Applied Linguistics*, 6, 132-146.
- Phillips, D. (1996). *Longman Preparation Course for the TOEFL TEST*. White Plains: Longman.
- Rost, M. (1990). *Listening in Language Learning*. New York: Longman Inc.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics* X, 209-30.
- Stern, H. H. (1975). What can we learn from the good language learner? *Canadian Modern Language Review*, 31, 304-318.
- Rost, M., & Ross, S. (1991). Learner use of strategies in interaction: typology and teachability. *Language Learning* 41: 2, 235-273.
- Rubin, J. (1975). What the good language learner can teach us. *TESOL Quarterly*, 9, 41-51.
- Varonis, E. M., & Gass, S. M. (1985). Non-native/non-native conversations: A model for negotiation of meaning. *Applied Linguistics*, 6, 71-90.