JALT Journal, Vol. 11, No. 2 (Nov. 1989)

Variation in the Relative Clause of Japanese Learners

Lynne Hansen-Strain
Brigham Young University, Hawaii
and
Jeris E. Strain
Himeji Dokkyo University

The research reported in this paper examines interlanguage variability in the comprehension and production of English relative clauses by Japanese second language learners.[1] Seven different instruments were used to elicit data from 75 students in a university English Language Institute in Hawaii, 15 from each of five L1 backgrounds: Japanese, Chinese, Korean, Samoan, and Tongan. Significant contextual variability was found in the relative clause performance of the Japanese students. While their performance was lower than other L1 groups on an aural comprehension test and an oral picture test, they performed more accurately than others on a written sentence combining task. The findings are interpreted as being supportive of a multiple competence model of SLA, and the implications of such a model for second language teaching and testing are discussed.

英語を学習する日本人の関係節修得の多様性について

この研究報告は、外国語として英語を学習する日本人の関係節の理解および使用についての多様性をのべたものである。この研究対象となったのは、ハワイにある英語教育研究所在籍の日本人・中国人・韓国人・サモア人・トンガ人の5か国語を母国語とする15人ずつ、合計75人であり、データの収集には7種類のテスト様式が用いられた。

日本人学習者の関係節使用能力については、そのコンテクストによる有意義な違いが見られた。聴解力テストおよび絵画口頭説明(oral picture)テストにおいては、他の母語者よりも低い成績を示したが、読解文結合(written sentence combining)テストでは、より正確な結果を示した。

この調査結果は第2言語修得の多様能力性の理論に合致するものであり、それに準拠した第2言語教育およびtestingの模範例の示すところを論じた報告書である。

1. Introduction

Grammatical accuracy in learner language has been found to vary according to the context of language use. That is, at any given stage of development a learner's interlanguage system appears to contain a number of competing rules, with one rule guiding performance on one occasion and another rule on a different occasion. Such contextual variability has been reported in L2 acquisition studies of phonology (Beebe, 1980; Dickerson, 1975; Sato, 1983) as well as of morphological and syntactic structures (Ellis, 1987a; Hansen-Strain, in press-a, in press-b; Larsen-Freeman, 1976; Tarone, 1985). This inherent variability of language learner behavior lies at the heart of basic assumptions in the second language acquisition field and is a fact of enormous importance for language teaching and testing. The theoretical position one chooses for explaining the nature and causes of variation has important consequences, not only for one's explanation of second language development, but for the practical concerns of syllabus design and the selection of teaching materials and language tests (Ellis, 1987b; Skehan, 1987).

The present paper examines the task-induced variability of one type of grammatical structure, the relative clause, in the interlanguage of one group of learners, adult Japanese ESL students. An analysis of the Japanese learner data is followed by statistical comparisons of these with data collected from learners from other first language backgrounds. Such cross-cultural research serves as a testing ground for the universal application of explanations of L2 variability, and offers insights into characteristics of the interlanguages of particular groups of learners.

2. Models of Interlanguage Variability

A number of different paradigms for viewing variation in interlanguage have been discussed in the SLA literature: homogeneous competence, dual competence, a capability continuum (Tarone, 1988), and multiple competence (Ellis, 1985). The homogeneous competence paradigm argues for a single dimension of language use, a unitary competence which guides language behaviour gen-

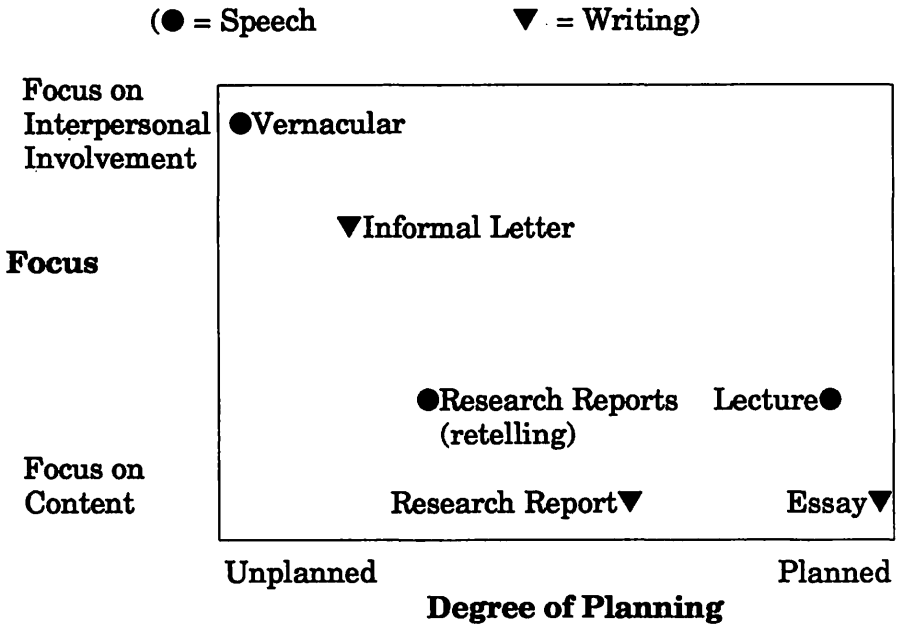
erally. The dual-competence paradigm of Krashen (1981) holds that performance based on acquisition and on learning will produce two different orders of accuracy for L2 structures deriving from an acquired competence which then may be augmented by conscious learning when conditions permit. The capability continuum paradigm of Tarone (1983) proposes that the capability underlying performance constitutes a range of styles along a continuum, and that this continuum is mainly affected by the amount of attention that is paid to speech. A multiple competence paradigm proposed by Ellis (1985) holds that interlanguage is composed of a series of variable systems which are domain specific. These are said to comprise a continuum of discourse types ranging from entirely unplanned to entirely planned.

The theoretical framework used in the present study is a multiple competence model of discourse variability which takes into account not only degree of planning but also language modality and level of interpersonal involvement (Hansen-Strain, in press-a). These parameters from orality/literacy studies (Chafe, 1982; Gee, 1986; Tannen, 1982) provide a framework for charting discourse types used within particular speech communities. In Figure 1, for example, we see several styles of English discourse represented. Within a given community, the available discourse types comprise the potential multiple competence of a language user in that community. Learning a second language involves the construction of these multiple competencies, the development of several interlanguages which are separate but overlapping in terms of role systems. The domain-specific interlanguages relate to the discourse types used by target language speakers. Through the use of this model, the types of discourse to be learned in a particular second language can be brought into focus, and compared with those in the learner's mother tongue. Such comparisons may well yield valuable insights leading to explanations for significantly different patterning of interlanguage variability between groups of L2 learners whose L1 discourse patterns differ significantly.

Evidence of the effects of first language background on task-induced variation is still quite limited. In some studies where variation is reported, the data from different L1 groups are not reported separately; in others, the subjects are drawn from a single L1 group. Tarone's 1985 study, however, provides information on

Figure 1

Discourse Variability



group differences in the use of four morphological and syntactic structures in English across three tasks. Although L1 differences in interlanguage were not a primary focus of her study, a research design providing for the separate reporting of data from the Asian and the Arabic subjects resulted in interesting evidence on group differences in variability. The elicitation tasks included a written grammar test, an oral interview, and an oral narrative. The most striking group difference in accuracy patterns was found in the production of the 3rd person singular *-s*. For the Japanese, there was no difference in accuracy for this morpheme over the three tasks, while the data from the Arabic students supported Tarone's hypothesis of decreasing accuracy as attention to form decreases (67% correct for the grammar test; 51% correct for the interview, and 39% for the narrative).

Table 1
Relative Clause Patterns in Six Languages

	English	Samoan	Tongan	Japanese	Chinese	Korean
Position with respect to head noun	Follow	Follow	Follow	Precede	Precede	Precede
Relative marker						
Obligatory	X*	—	—	—	X	X
Optional	—	X	X	—	—	—
Not used	—	—	—	X	—	—
Relative marker morphology	Variant	Variant	Invariant	—	Invariant	Invariant
Positions relativizable						
Subject	X	X	X	X	X	X
Direct Object	X	X	X	X	X	X
Indirect Object	X	X	X	X	X	X
Object of Preposition	X	X	X	X	X	X
Genitive	X	X	X	X	X	X
Object of Comparative	X	X	X	—	—	—
Pronoun Retention						
Subject	—	—	—	—	—	—
Direct Object	—	X	X	—	—	—
Indirect Object	—	X	X	—	—	—
Object of Preposition	—	X	X	—	X	—
Genitive	—	X	X	X**	X	X**
Object of Comparative	—	X	X	—	—	—

*Marker is optional when it is not a subject and not initial in its clause.
**Retention is optional in this position.

3. Relative Clause Acquisition

The relative clause, widely investigated in SLA research, is a structure which varies in a number of respects among the languages of the world. Some of the parameters of this variation are illustrated in Table 1, a summary of characteristics of relativization in the six languages included in the present study. Here we see

Table 2

Pronominal Copy Retention

Subject:	the boy that *he* came
Direct Object:	the boy that John hit *him*
Indirect Object:	the boy that I sent a letter to *him*
Object of Preposition:	the boy that I sat near *him*
Genitive:	the boy that *his** father died
Object of Comparative:	the boy that John is taller than *him*

*Note retention of genitive copy *his* precludes "whose."
Examples from Schachter (1974).

that relative clauses may precede the head noun (as in Japanese, Chinese and Korean) or may follow the head noun (as in English, Samoan, and Tongan). The relative marker may not be present at all (as in Japanese), may be obligatory (as in English, Chinese, and Korean), or may be optional (as in Samoan and Tongan). In languages where a relative marker is used, its morphology may be variant (as in English, and Samoan) or invariant as in (Tongan, Chinese, and Korean).

Languages also differ with regard to the positions which can be relativized. Little variation is seen along this parameter among the six languages included in Table 1 since they are relativizable in all six positions (Subject, Direct Object, etc.), with the exception of Object of Comparative for the three Asian languages. The languages exhibit great differences, however, in the relative clause positions which require pronoun retention, that is, retention of pronominal copies. These, as illustrated in Table 2, are inappropriate in all positions in English. In contrast, some languages of the world, such as the Polynesian languages in the present study, require pronoun retention in most positions.

The patterning of occurrence of positions that can be relativized and of positions requiring pronominal copies in the languages of the world is accounted for by a universal hierarchy of grammatical relations hypothesized by Keenan and Comrie (1977). Based on an investigation of relative clause formation strategies in a broad range of languages, this Accessibility Hierarchy (AH) suggests that there is a universal order of grammatical relations out of which

relativization can take place:

SU > DO > IO > PREP > GEN > COMP

For example:

SU (Subject)	the boy that came
DO (Direct Obj)	the boy that John hit
IO (Indirect Obj)	the boy that he spoke to
PREP (Obj of Prep)	the boy that he sat near
GEN (Genitive)	the boy whose father died
COMP (Obj of Comp)	the boy that he is taller than

This hierarchy can be interpreted as an implicational scale of markedness with the *subject* and the *object of the comparative* being respectively the least and most marked positions on the scale. If, in a given language, a relative clause can be formed with the relativizable noun phrase in a certain position, then it will also be the case that in that language one can form relative clauses with the coreferential noun phrase bearing any grammatical relation listed to the left of that particular position on the hierarchy.

In the application of AH to second language acquisition, as the basis for the prediction of accuracy and acquisition orders, relative clauses formed on the subject are predicted to be easiest to learn, those on the object of a comparative most difficult. These predictions based on AH have been confirmed in large measure by studies examining the L2 acquisition of relative clauses in English, but not for every detail in every case. For example, studies done by Pavesi (1986) and Gass (1979) found opposite patterns of acquisition with regard to the genitive position on AH. The L2 English data that were collected from Pavesi's (1986) two groups of L1 Italian subjects (EFL high school students in Italy and ESL migrant workers in Great Britain) followed the AH. (These data were elicited orally using the Hyltenstam (1984) picture test, an instrument also used in the present study). Gass (1979), on the other hand, using data collected from foreign students of various L1 backgrounds who were enrolled in a university ESL program in the United States, found more correct responses for the genitive position than would have been predicted by AH. On both sentence combining and grammatical judgment tasks (instruments also used in this study), Gass found an exception to AH predictions:

rather than maintaining its low position on the hierarchy, the genitive was second only to the subject in number of correct responses of the learners.

Keenan and Comrie's Accessibility Hierarchy has also been used as a hypothesis in accounting for the occurrence of pronominal copy retention (see Table 2) in learner language. Pronominal copies have been found in the interlanguages of second language learners from a number of language backgrounds, whether or not their L1s allowed it. Ioup and Kruse (1977) and Gass (1979) found this to be the case for L2 English, as did Hyltenstam (1984) for L2 Swedish. Further, the frequency of the copies almost always followed Keenan and Comrie's hierarchy. A comparison of copy frequencies for the Swedish learners across L1 groups, however, did indicate apparent first language influence in the *strength* and *duration* of the learners' pronominal retention strategy.

First language influence in the L2 English relative clause was also reported by Schachter (1974), who found L1 transfer to manifest itself in terms of the number of relative clauses used in the written themes of her adult ESL students, though not in terms of the number of errors that were made. The Japanese and Chinese subjects, for whom English relative constructions were predicted to be most difficult (on the basis of contrastive analyses of these languages with English), used relativized constructions far less frequently than the Arabic and Persian students in the study. Kleinmann (1978) also found that students from different first language backgrounds (Arabic, Spanish/Portuguese) showed differences in the frequency of relative clause usage.

A number of investigations of L2 relative clause acquisition, however, have found no evidence of L1 differences. Cook (1973), for example, tested comprehension by means of imitation tasks given to children acquiring English as a first language and to adult second language learners. Since the error types for both groups are similar, Cook concluded that native language transfer does not play a significant role in learning. The same conclusion was reached by Ioup and Kruse (1977) who elicited grammaticality judgments from university ESL learners and found nonsignificant differences between groups based on language background.

In the L2 research literature reviewed here, we see quite divergent views as to the influence of first language background on

relative clause acquisition. The contradictory findings may not seem so surprising, though, considering the broad array of different data collection procedures that were used in these studies: grammatical judgments, picture elicitation, essays, imitation tasks, and sentence combining. From the perspective of a multicompetence model of interlanguage, these different methods of data collection could be seen as accessing separate interlanguage competencies, each by itself providing an incomplete picture of the language that had been learned. The present investigation differs from previous ones in that a more complete picture is elicited through the use of seven different instruments administered to the same subjects. These include production as well as comprehension tasks, both oral and written. The research questions posed are the following:

1. What is the patterning of the Japanese performance profile for the relative clause eliciation tasks?
2. Does the relative clause performance of the Japanese learners differ significantly from that of other L1 groups in:
   a) performance profiles over the various elicitation tasks?
   b) the extent to which L2 relative clause comprehension and production follow Accessibility Hierarchy predictions?
   c) pronominal copy errors in L2 relative clause production?

4. Method

4.1 Subjects.

Fifteen Japanese ESL learners participated in the study, together with 15 learners from each of four other L1 backgrounds: Chinese, Korean, Samoan, and Tongan. The subjects were randomly selected from ESL reading classes in the English Language Institute at Brigham Young University–Hawaii. The 38 male and 37 female students were from the four proficiency levels at the institute, with scores ranging between 46 and 84 on the Michigan Test of English Language Proficiency (MTELP), the external criterion used for this study. They ranged in age from 17 to 29. Table 3 is a breakdown by culture of means for MTELP scores. A one-way ANOVA (analysis of variance of means) by culture for the five groups indicates no significant difference between them in

Table 3

MTELP Scores by Culture

Culture	MTELP Mean	S.D.	N
Samoan	69	8.5	15
Tongan	68	10.3	15
Japanese	67	9.1	15
Chinese	72	7.2	15
Korean	68	13.3	15
(*p*	n.s.	n.a.	n.a.)

their English proficiency level as measured by the Michigan Test. In addition to the 75 second language learners, 15 native speakers of standard American English, university students between the ages of 17 and 21, were included in the study as an L1 comparison group for five of the instruments.

4.2 Materials and procedures.

Relative clause data were collected using seven different instruments: 1) aural comprehension; 2) picture test; 3) oral retelling; 4) written retelling; 5) essay; 6) sentence combining; and 7) grammatical judgments. Of these 1, 2, and 3 are oral/aural tasks; 4, 5, 6, and 7 are written.

The *aural comprehension test* consists of two sets of recorded sentences, each set containing two of each of nine relative clause sentence types (SS, SO, SI; OS, OO, OI; IS, IO, II; e.g., SS: *the cow that hit* the dog *(V)* (O); OS: (S) *(V) the horse that hit* the dog; IS: (S) *(V) to the cow that hit* the horse; etc.). Subjects were tested individually and assigned randomly to either set. The tester was a senior TESL major at BYUH, a 23-year-old native speaker of English from Alaska. She required the students to act out the sentences they heard using toy animals, following Sheldon's (1974) procedure. Four different animals were used: "cow," "horse," "dog," and "pig." The verbs used were "bumped," "hit," "killed," and "pushed" (taking direct objects) and "shouted to" and "whispered to" (taking indirect objects). The taped sentences were each presented twice with a subsequent 15-second pause. As the subject

manipulated the animals, the investigator recorded the responses on a coding sheet.

The *picture test* is the instrument developed by Hyltenstam (1984) for his study of the acquisition of Swedish relative clause structures by second language learners. The material consists of six posters, one for each NP position on Keenan and Comrie's (1977) Accessibility Hierarchy (see above). On each poster there are eight pictures: two each of men, women, boys, and girls; two verbs are portrayed on each poster. The eight pictures are numbered 1 to 8, and the subjects' task is to orally identify the person in each numbered picture in answer to the question, "Who is number X?" The subjects were tested individually as follows: first five responses were elicited for each poster, and then one response each for the six posters. The interviews, which lasted from about 10 to 25 minutes, were audio-taped and later coded by the experimenter, the same native-speaking research assistant who administered the aural comprehension test.

The *oral and written retellings* were of a 500-word passage, a summary of Suter's (1976) study of "Factors in Pronunciation Achievement." After subjects had been asked to study the passage for fifteen minutes, the reading and any notes they had made were collected by the examiner, a teacher in the English Language Institute. For the written retelling, the students were then given twenty minutes to write what they had read; and for the oral retelling, twenty minutes to retell it into a cassette recorder.

The *sentence combining* instrument is the one used by Gass (1979). Subjects were presented with twelve pairs of sentences and were instructed to combine each pair to form one English sentence. Each pair of sentences represented one step on the Accessibility Hierarchy, with six items in pre-verb (subject) position and six in post-verb (object) position. The instructions attempted to preclude the writing of sentences other than those containing relative clauses.

The *grammatical judgments* task (Gass 1979) requires judgments of 29 sentences of which 13 are grammatical and 16 ungrammatical. The ungrammatical sentences represent four error types: (a) relative clause marker omission; (b) pronoun retention; (c) relative clause marker morphology; and (d) adjacency. After reading

each of the sentences, the subjects simply indicated whether they considered it to be grammatical or ungrammatical.

All L2 data were collected during the fall 1986 semester in the English Language Institute. The written retellings were collected in September, and the picture elicitation, oral comprehension, grammatical judgments, and sentence combining data during October and November. The oral retellings were collected in December, as were the essays, which were part of a final examination test battery.

5. Results

5.1 Japanese performance profile

In order to facilitate the comparison of relative clause performance across elicitation tasks, the accuracy scores (for aural comprehension (AC), picture test (PT), sentence combining (SC), and grammatical judgments (GJ)) and frequency totals (for oral retelling (OR), written retelling (WR), and essay (R)) were calculated and converted to z-scores (i.e., the percentage means and numerical means of the two types of instruments were converted to their standard deviations above (+) and below (-) their individual means). Table 4 shows the distribution of these z-scores for each of the seven instruments by culture. The three tests on the left of the chart are in the spoken modality; the four on the right the written. The Degree of Planning continuum (see Figure 1) extends, on the left, from least opportunity for planning to most planning opportunity on the right. The continuum for Focus begins at the top with most interpersonal involvement (spontaneous) and extends to complete concern with content at the bottom. Figure 2 represents the L2 data from Table 4 in graph form.

The patterning of variability in the Japanese performance profile shows that accuracy is lowest for the instruments on the left side of the chart, the oral/aural tests which allow least opportunity for planning and most personal involvement. In other words, the competence of the Japanese subjects to comprehend spoken sentences containing relative clauses and to use them in spoken discourse appears low when compared with the abilities they demonstrate in using the same constructions in writing.

Table 4

Z-Scores* for Seven Tasks by Culture

Culture	AC	PT	OR	WR	E	SC	GJ
Samoan	.918	.577	.432	.000	.268	-.125	-.021
Tongan	.737	.122	.199	-.440	.354	-.379	-.008
Japanese	-.492	-.401	-.097	.250	-.073	.557	-.046
Chinese	-.703	-.054	-.052	.059	-.207	-.148	.038
Korean	-.360	-.237	-.554	-.059	-.304	.066	.037
($p <$	.001	.05	.05	n.s.	n.s.	n.s.	n.s.)

*+/- standard deviations for percentage means (AC, PT, SC, GJ) and numerical means (OR, WR, E).

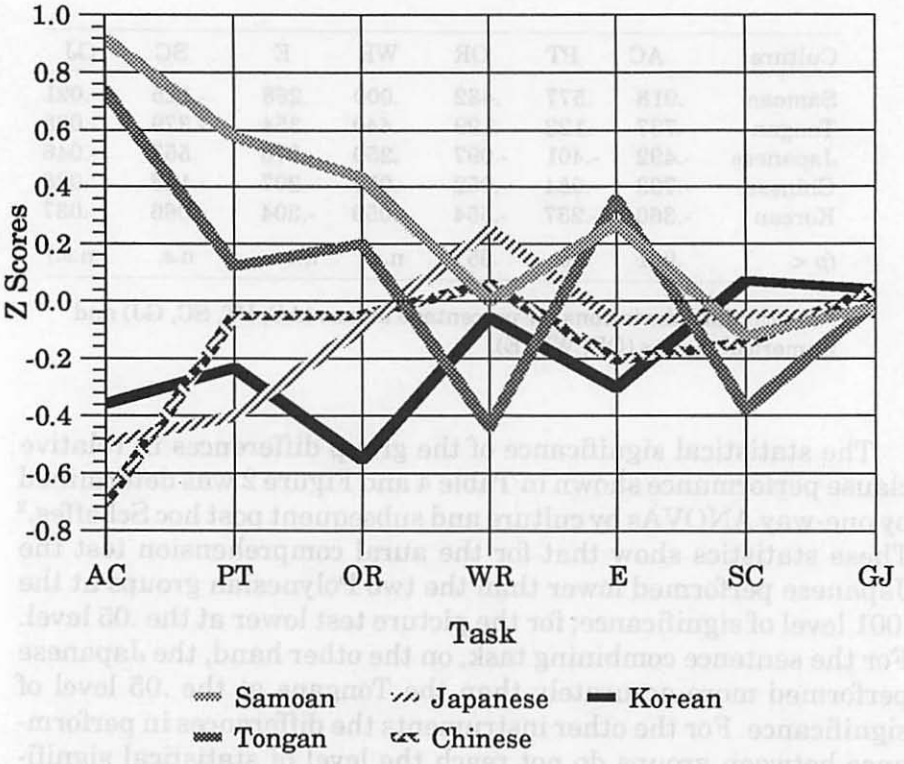
The statistical significance of the group differences in relative clause performance shown in Table 4 and Figure 2 was determined by one-way ANOVAs by culture and subsequent post hoc Scheffes.[2] These statistics show that for the aural comprehension test the Japanese performed lower than the two Polynesian groups at the .001 level of significance; for the picture test lower at the .05 level. For the sentence combining task, on the other hand, the Japanese performed more accurately than the Tongans at the .05 level of significance. For the other instruments the differences in performance between groups do not reach the level of statistical significance.

5.2 Comparison of performance profiles

Turning to the research questions, we find that although the five groups of ESL learners are at the same level of English proficiency (as determined by the ANOVA of MTELP scores by culture), there are highly significant group differences in performance on three of the elicitation tests. On the one hand, the performance of the Japanese students falls below that of the Polynesians on the oral/aural tasks which offer minimal opportunity for planning and a maximum of personal involvement (aural comprehension and picture test); on the other hand, the Japanese students' accuracy exceeds that of the Polynesians on a written task with maximum

Figure 2

Relative Clause Z-Scores for Seven Tasks by Culture



opportunity for planning and a minimum of personal involvement. In general these findings suggest that the Japanese students comprehend relative clause constructions less well than the Polynesians in the spoken modality (aural comprehension test), and that they form sentences containing relative clauses less accurately in spoken English (picture test) than they do in written English (sentence combining).

As shown in Table 5, the aural comprehension instrument elicited the most striking group differences in performance. An item analysis given in Table 6 shows that for eight of the nine sentence types on the test the Tongans' and the Samoans' comprehension far exceeded that of the Japanese and other Asian groups.

Table 5

Aural Comprehension Test Descriptive Statistics by Culture

Culture	Mean	S.D.	N
Samoan	15.3	1.4	14
Tongan	14.3	3.2	15
Japanese	7.7	4.7	15
Chinese	6.6	4.8	14
Korean	8.4	5.1	14
L1 English	17.2	1.6	15
(L2 groups: $p < .001$)			

The exception is the SS type, the most accessible of all in terms of the AH hypothesis; it was understood well by all.

It is interesting to compare these results from the aural comprehension test, on one end of the planning and interpersonal involvement continua, to those from the task on the other end, the grammatical judgments test. Table 7 gives the descriptive statistics for this instrument, and Table 8 a breakdown of performance for each sentence type. In sharp contrast with the aural comprehension test, the differences between the L1 groups in performance on the grammatical judgments were not significant for the test as a whole, nor for any of the positions on the hierarchy.

Table 6

Aural Comprehension Mean Test Scores for the Nine Sentence Types by Culture

Culture	SS	SO	SI	OS	OO	OI	IS	IO	II
Samoan	1.5	1.2	1.5	1.7	1.6	1.7	1.6	1.6	1.9
Tongan	1.7	1.3	1.2	1.6	0.2	1.7	1.7	1.7	1.6
Japanese	1.7	0.7	0.6	0.5	0.8	1.1	0.6	0.5	1.0
Chinese	1.6	0.4	0.5	0.4	0.6	1.1	0.4	0.6	0.9
Korean	1.5	0.6	0.8	0.7	0.8	1.2	0.9	0.9	1.1
L1 English	2.0	1.8	1.9	1.9	2.0	1.9	1.9	1.9	1.8

(ANOVAs were not run.)

Table 7

Grammatical Judgments Descriptive Statistics by Culture

Culture	Mean	S.D.	N
Samoan	33.4	2.8	15
Tongan	34.1	3.1	15
Japanese	32.6	4.1	15
Chinese	35.6	3.4	15
Korean	35.6	4.4	15
L1 English	45.1	1.2	15

(L2 groups: $p <$ n.s. for all categories)

Although significant group differences in relative clause performance were not found for the paper-pencil tests (written retelling, essay, sentence combining, and grammatical judgments), the significant differences in performance on the other tasks (aural comprehension, picture test, and oral retelling) do warrant the conclusion that the Japanese performance profile does differ significantly from the other L1 groups, and (as seen in Table 4 and Figure 2) most sharply with the two Polynesian groups. The Japanese subjects appear to know English relative clauses better when they are tested using written tests, which allow more plan-

Table 8

Percentage Correct on Grammatical Judgments Task for AH Positions by Culture

LI Group	SU	DO	IO	GEN	COMP
Samoan	78	82	60	70	43
Tongan	74	80	53	70	43
Japanese	78	72	57	83	53
Chinese	65	68	40	80	36
Korean	67	60	51	70	40
L1 English	88	91	88	98	86

(L2 groups: $p <$ n.s. for all categories)

Table 9

Sentence Combining Test Mean Scores
for Relative Clause Types by Culture

LI Group	SU	DO	IO	Prep	Gen	Comp
Samoan	1.66	1.53	1.00	.40	.80	.73
Tongan	1.43	1.29	.71	.43	.78	.71
Japanese	1.47	1.66	1.73	.73	1.60	1.00
Chinese	1.40	1.20	1.00	.40	1.33	.73
Korean	1.48	1.21	1.14	.71	1.28	.93
($p <$	n.s.	n.s.	n.s.	n.s.	.05	n.s.)

Picture Test Mean Scores for Relative Clause Types by Culture

LI Group	SU	DO	IO	Prep	Gen	Comp
Samoan	5.0	4.6	5.3	4.7	4.8	5.3
Tongan	5.7	4.9	4.7	4.7	4.6	4.0
Japanese	5.6	4.5	3.5	3.2	4.9	3.7
Chinese	5.7	5.8	4.1	3.3	5.1	3.6
Korean	5.6	4.4	3.6	3.3	5.3	3.6
($p <$ n.s. for all categories)						

ning and less interpersonal involvement. For the Polynesians, on the other hand, the relative clause acquisition level appears to be higher when aural/oral elicitation tasks with minimum planning and maximum personal involvement are used.

5.3 Accessibility Hierarchy

In answering the question concerning group differences in the predictive accuracy of the Accessibility Hierarchy, we turn to analyses of responses given on the four instruments which specifically elicit relative clause performance on the AH positions: the aural comprehension and grammatical judgments task data presented above, together with the picture test and sentence combining task data. Table 9 gives the mean group scores on the latter two tests for the six sentence types on the hierarchy. These figures

indicate that the subjects in our study do follow the predictions of AH, with one notable exception, the genitive position. Like Gass' (1979) subjects (but unlike those of Pavesi, 1986), the ESL learners in the present study find relative clauses formed on genitives to be considerably easier than their low place on the hierarchy would predict.

Group differences on the genitive position are seen in comparing the L1 data, however. Three of the elicitation instruments are relevant for the comparison: picture test, sentence combining, and grammatical judgments (the aural comprehension test cannot provide evidence on the genitive since it elicited only the first three positions on the hierarchy). On the picture elicitation the performance of the Japanese students and that of the other Asian groups follows the predictions of the accessibility hierarchy, with the exception of the genitive. Both of the Polynesian groups, on the other hand, have high scores for all test items, with slight differences in performance between the six sentence types. On the sentence combining test, again we see high scores on the genitive for the Asian groups, with the Japanese differing from the Polynesians on this position statistically, at the .05 level, as indicated by a one-way ANOVA by culture and post hoc Scheffe analysis. Although the South Pacific students also score higher on the genitive than predicted by AH, still their control of relative clauses formed on this position appears to be at a lower level relative to other positions on the hierarchy than their Asian classmates. The performance on the grammatical judgments test (Table 7) shows the most uniformity of the three tests among the L1 groups. In fact, no significant group differences are found for any of the sentence types (although we notice again a trend for higher scores on the genitive for Japanese and Chinese than for Polynesians), and again clear evidence of higher scores overall for clauses formed on genitives than predicted by AH. In comparing group performance across these three tasks then, we see that the facilitation of the genitive position, although present across all five groups of learners, is more pronounced for the Japanese subjects than for the Polynesians.

A further source of evidence on group differences in the applicability of Keenan and Comrie's AH to learner language is the *frequency* of the six relative clause types in L2 spoken and written discourse. Such frequencies are presented in Table 10 for the oral

Table 10

Picture Test Mean Scores for Relative Clause Types by Culture

Oral Retelling

LI Group	SU	DO	IO	Prep	Gen	Comp
Samoan	4.6	2.5	0.2	1.1	0.2	0.0
Tongan	3.4	3.1	0.0	0.5	0.2	0.0
Japanese	3.8	1.1	0.0	0.2	0.2	0.0
Chinese	3.7	1.3	0.0	0.2	0.1	0.0
Korean	2.2	1.1	0.0	0.1	0.1	0.0
American	4.4	2.0	0.1	0.2	0.3	0.0

Written Retelling

LI Group	SU	DO	IO	Prep	Gen	Comp
Samoan	1.1	0.2	0.0	0.2	0.0	0.0
Tongan	0.5	0.2	0.0	0.2	0.0	0.0
Japanese	1.2	0.2	0.0	0.0	0.1	0.0
Chinese	1.5	0.3	0.0	0.0	0.0	0.0
Korean	1.4	0.10	0.0	0.0	0.0	0.0
American	1.2	0.5	0.0	0.3	0.1	0.0

Essay

LI Group	SU	DO	IO	Prep	Gen	Comp
Samoan	2.9	2.9	0.0	0.8	0.0	0.0
Tongan	4.0	2.7	0.0	0.5	0.0	0.0
Japanese	2.6	1.8	0.1	0.4	0.0	0.0
Chinese	1.6	1.4	0.0	0.5	0.0	0.0
Korean	1.2	1.3	0.0	0.1	0.0	0.0
American	3.1	2.2	0.0	0.4	0.0	0.0

(ANOVAs were not run.)

Table 11

Z-Scores for Copy Retention Errors by Culture

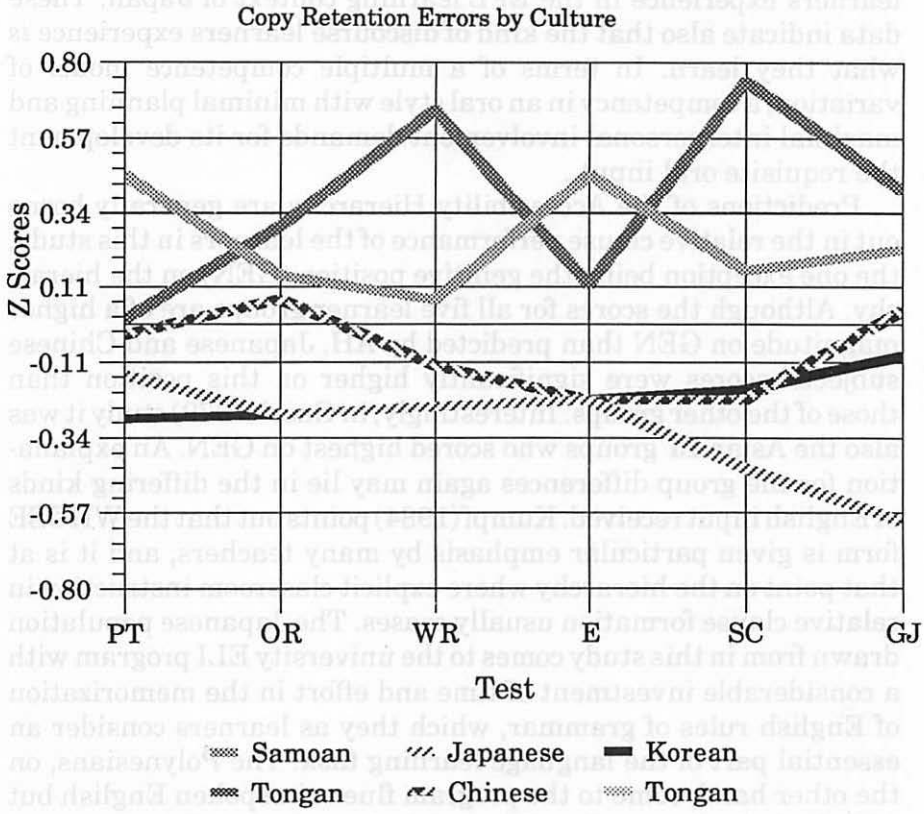
Culture	PT	OR	WR	E	SC	GJ
Samoan	.458	.150	.072	.459	.163	.230
Tongan	.013	.300	.655	.122	.742	.411
Japanese	-.156	-.275	-.250	-.230	-.437	-.600
Chinese	-.035	.075	-.127	-.230	-.228	.041
Korean	-.287	-.275	-.250	-.230	-.197	-.096
($p <$	n.s.	n.s.	n.s.	n.s	0.1	0.1)

and written retellings and the essay. The means in the table represent the mean number of relative clauses appearing per subject in the data from each L1 group. Notice that for each of the three tasks and for all of the L1 groups the pattern of use of the relative constructions does generally follow predictions of the hierarchy. Relatives formed on subjects are by far the most frequent, with those on direct objects next. Slight deviations from the predicted frequencies appear on the lower positions which are all characterized by low numbers of occurrence. For all three tasks more relatives are formed on objects of prepositions than on indirect objects, counter to AH predictions, and a trend can be noted for the Polynesians to relativize on the Prep position more frequently than the Asians do.

5.4 Copy retention

Finally, turning to our question concerning the incidence of copy retention errors on the elicitation tasks, Table 11 and Figure 3 summarize the z-scores for pronominal copy frequency. The mean z-scores represent the total number of pronominal or nominal copies used by each L1 group, so the larger the number on the table and the higher the line on the chart, the more copies were used in the interlanguage of that group. Notice that in most cases it is the Japanese students who make the fewest copy errors. For all of the tests, it is the South Pacific students who make more copy errors than do the Asian groups, just as contrastive analyses of their

Figure 3



respective first languages with English would predict. For the sentence combining and the grammatical judgments the group differences were significant, at the .01 level, as determined by one-way ANOVAs by culture.

6. Discussion

The present study shows contextual variability in the relative clause performance of Japanese learners of English. Their performance profile (Figure 2) for the elicitation tasks shows lower accuracy on listening comprehension and speaking tests than on written ones. A factor in the patterning of this variation may be the

limited exposure to spoken English discourse which Japanese learners experience in the EFL learning context of Japan. These data indicate also that the kind of discourse learners experience *is* what they learn. In terms of a multiple competence model of variation, a competency in an oral style with minimal planning and maximal interpersonal involvement demands for its development the requisite oral input.

Predictions of the Accessibility Hierarchy are generally borne out in the relative clause performance of the learners in this study, the one exception being the genitive position (GEN) on the hierarchy. Although the scores for all five learner groups are of a higher magnitude on GEN than predicted by AH, Japanese and Chinese subjects' scores were significantly higher on this position than those of the other groups. Interestingly, in Gass' (1979) study it was also the Asian L1 groups who scored highest on GEN. An explanation for the group differences again may lie in the differing kinds of English input received. Kumpf (1984) points out that the WHOSE form is given particular emphasis by many teachers, and it is at that point on the hierarchy where explicit classroom instruction in relative clause formation usually ceases. The Japanese population drawn from in this study comes to the university ELI program with a considerable investment of time and effort in the memorization of English rules of grammar, which they as learners consider an essential part of the language learning task. The Polynesians, on the other hand, come to the program fluent in spoken English but with little background or interest in learning explicit grammar rules. It may be the greater emphasis on grammatical form in Asian instructional settings, in this case a greater emphasis on the structure of relative clauses formed on GEN, that contributes to the Japanese students' relatively high performance on them.

The low incidence of pronominal copy errors in the Japanese data in comparison with the Polynesian appears to stem from structural properties of the mother tongues. Contrastive analyses would predict, due to the lower occurrence of copies in Japanese than in Samoan and Tongan, that the Japanese would indeed commit fewer copy retention errors in English than their Polynesian classmates. Furthermore, principles of Universal Grammar absolutely predict that since the presence of pronominal copies in a

language is marked, if found in a learner's L1, copy retention will be carried over into the L2 (White 1987).

Summarizing, then, the answers to the questions posed at the beginning of this paper, we have found that Japanese learners of English form relative clauses more accurately and understand them better in the written modality than the spoken. In comparison with L1 groups of the same English proficiency level from oral cultures in the South Pacific, they (a) performed significantly lower on the aural comprehension test ($p > .0001$) and on the oral picture test ($p > .05$); (b) formed genitive relative clauses with greater accuracy in comparison with their accuracy on other positions on the Accessibility Hierarchy; and (c) made significantly fewer pronominal copy errors.

The findings of the study are compatible with a multiple competence model of SLA. They suggest not only that interlanguage varies systematically across elicitation tasks, but also that the patterning of this variability differs significantly between groups from different first language backgrounds. A homogeneous competence model appears to offer little in the way of explanation for these phenomena. Viewed within the framework of a multiple competence paradigm of variation, however, these data can be seen as deriving from separate interlanguage competences which develop uniquely for each group of learners according to the features of L2 (and possibly L1) discourse experienced in the native culture, both in and out of the classroom.

These findings have important implications for language test selection and interpretation. As Tarone (1985) has pointed out, most teachers assume that learners' speech performance will probably be less accurate than their performance on a paper-and-pencil classroom test. The results presented here, however, support Tarone's observation based on her own data that "it may be that when ESL students argue that classroom tests do not really measure their ability to speak English grammatically, they are right" (p. 386).

Based on the findings of the present study, we could add that learners from some L1 backgrounds are "more right" than others in making this assessment of the testing of their second language proficiency. The discrepancy between accuracy in performance on oral

and written tasks tends to be greater for some groups of learners than for others and, in fact, may be in opposite directions for different groups. For example, the Polynesians' comparatively high level of relative clause mastery as indicated on aural/oral tests is belied by their lower scores on written tasks. Their Asian classmates, on the other hand, generally appear to have lower relative clause competence on the aural/oral tests than they do on the written.

Thus, in addition to the obvious general implications of contextual variability research for language testing, the findings of the present study have particular applications to test selection and interpretation in multicultural contexts. Social and ethical considerations in such settings favor the development of proficiency profiles based on a number of different tests (Evans and Hansen-Strain, 1986; Spolsky, 1987). The use of several proficiency measures in identifying such profiles can help to uncover group biases or performance tendencies which may be inherent in particular testing techniques.

In affirming an SLA model which recognizes language variability, the present research also recommends approaches to language teaching which are based on the acquisition of multiple competences in the target language. In an important paper which examines second language teaching within a framework of interlanguage variability, Ellis (1987b) stresses the dual contribution of formal as well as informal instruction; of accuracy as well as fluency. He suggests that these can be fostered by a parallel syllabus incorporating both product and process elements leading to materials that contain a broad range of discourse types. The assumption is that in classroom language learning the development of multiple competencies rests on the availability of opportunities to take part in different kinds of classroom interaction.

The variability patterns reported in this study suggest that the emphasis in Japanese EFL classrooms is on planned, written styles with little interpersonal involvement. As teachers we need to be aware that learning to perform in such careful styles is quite different from learning how to perform in unplanned spoken ones. As Ellis (1987b) comments, "Because different kinds of knowledge and different processes of language use are involved in different discourse types, it cannot be expected that the acquisition of one

style will facilitate the use of another style" (p. 192). Thus a primary concern of a second language teacher is to insure a match between the interactional opportunities available to the learner and the kind of competence the teaching is designed to produce.

Lynne Hansen-Strain is a professor of linguistics at Brigham Young University-Hawaii where she teaches in the English Language Institute and in the B.A.-TESL program. Previously she taught English in Nigeria, Austria, and Pakistan. Her current research interests are in individual and group differences in second language acquisition, particularly in the acquisition of Japanese as a second language.

Jeris E. Strain is a professor of English at Himeji Dokkyo University. Previously he taught English in Michigan, Iran, Egypt, Bulgaria, and Hawaii. His current research interests are in ESL/EFL methodology. His most recent publication is *Charles Carpenter Fries: His Oral Approach* (with Wm. Norris).

Notes

1. The research presented in this paper was supported by a grant from the Institute for Polynesian Studies, Brigham Young University—Hawaii, for which we are grateful. In addition, thanks are due to Clyde Robinson for assistance with the statistical analyses, Song Xing-Jun for help with the coding and data entry, and Teri Lee Lehman for assistance with data collection; also, for their cooperation as we collected data, to teachers in the English Language Institute, Norm Evans, Mark James, Uinise Langi, and Earl Wyman.
2. The Scheffes procedure is used to identify where differences occur for means shown by the ANOVAs to be significant.

References

Beebe, L. (1980). Sociolinguistic variation and style shifting in second language acquisition. *Language Learning, 30*, 433-447.

Chafe, W. (1982). Integration and involvement in speaking, writing, and oral literature. In D. Tannen (Ed.), *Spoken and written language: Exploring orality and literacy* (pp. 35-53). Norwood, NJ: Ablex.

Cook, V. (1973). The comparison of language development in native children and foreign adults. *IRAL ,11*, 13-28.

Dickerson, L. (1975). Interlanguage as a system of variable rules. *TESOL Quarterly, 9*, 401-407.

Ellis, R. (1985). *Understanding second language acquisition.* Oxford: Oxford University Press.

Ellis, R. (1987a). Interlanguage variability in narrative discourse. *Studies in Second Language Acquisition, 9,* 1-20.

Ellis, R. (1987b). Contextual variability in second language acquisition and the relevancy of language teaching. In R. Ellis (Ed.), *Second language acquisition in context* (pp. 179-194). Englewood Cliffs, NJ: Prentice Hall International.

Gass, S. (1979). *An investigation of syntactic transfer in adult second language acquisition.* Unpublished doctoral dissertation, Indiana University.

Gee, P. G. (1986). Orality and literacy: From the savage mind to ways with words. *TESOL Quarterly, 20,* 719-746.

Hansen-Strain, L. (in press-a). Typological universals and relative clause in interlanguage. In H. Nicholas (Ed.), *Explaining interlanguage development.* Clevedon, Avon: Multilingual Matters.

Hansen-Strain, L. (in press-b). Orality/literacy and group differences in second language acquisition. *Language Learning.*

Hyltenstam, K. (1984). The use of typological markedness conditions as predictors in second language acquisition: The case of pronominal copies in relative clauses. In R. Anderson (Ed.), *Second languages: Across-linguistic perspective* (pp. 39-58). Rowley, MA: Newbury House.

Ioup, G., & Kruse, A. (1977). Interference versus structural complexity as a predictor of second language relative clause acquisition. In C. Henning (Ed.), *Proceedings of the second language acquisition research forum* (pp. 48-60). Los Angeles: UCLA.

Keenan, E., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry, 8,* 63-100.

Kleinmann, H. (1978). The strategy of avoidance in adult second language acquisition. In W. C. Ritchie (Ed.), *Second language acquisition research: Issues and implications* (pp. 157-174). New York: Academic Press.

Krashen, S. (1981). *Second language acquisition and second language learning.* Oxford: Pergamon Press.

Kumpf, L. (1984). Comments on the paper by Gass and Ard. In W. Rutherford (Ed.), *Language universals and second language acquisition* (pp. 69-72). Amsterdam: John Benjamins.

Larsen-Freeman, D. (1976). An explanation for the morpheme acquisition order of second language learners. *Language Learning, 26,* 125-134.

Pavesi, M. (1986). Markedness, discoursal modes and relative clause formation in a formal and an informal context. *Studies in Second Language Acquisition, 8,* 38-55.

Sato, C. (1983). Task variation in interlanguage phonology. Paper presented at the 10th University of Michigan Conference on Applied Linguistics, Ann Arbor.

Schachter, J. (1974). An error in error analysis. *Language Learning, 24,* 205-214.

Sheldon, A. (1974). On the role of parallel function in the acquisition of relative clauses in English. *Journal of Verbal Learning and Verbal Behavior, 13,* 272-281.

Skehan, P. (1987). Variability and language testing. In R. Ellis (Ed.), *Second*

language acquisition in context (pp. 195-206). Englewood Cliffs, NJ: Prentice Hall International.

Spolsky, B. (1987). Social aspects of individual assessment. Paper presented at the 8th World Congress of Applied Linguistics, Sydney, Australia.

Suter, R. W. (1976). Predictors of pronunciation accuracy in second language learning. *Language Learning, 26*, 233-253.

Tannen, D. (1982). The oral/literate continuum in discourse. In D. Tannen (Ed.), *Spoken and written language: Exploring orality and literacy* (pp. 3-17). Norwood, NJ: Ablex.

Tarone, E. (1983). On the variability of interlanguage systems. *Applied Linguistics, 4*, 142-163.

Tarone, E. (1985). Variability in interlanguage use: A study of style shifting in morphology and syntax. *Language Learning, 35*, 373-403.

Tarone E. (1988). *Variation in interlanguage.* London: Edward Arnold.

White, L. (1987). Markedness and second language acquisition: The question of transfer. *Studies in Second Language Acquisition, 9*, 261-286.