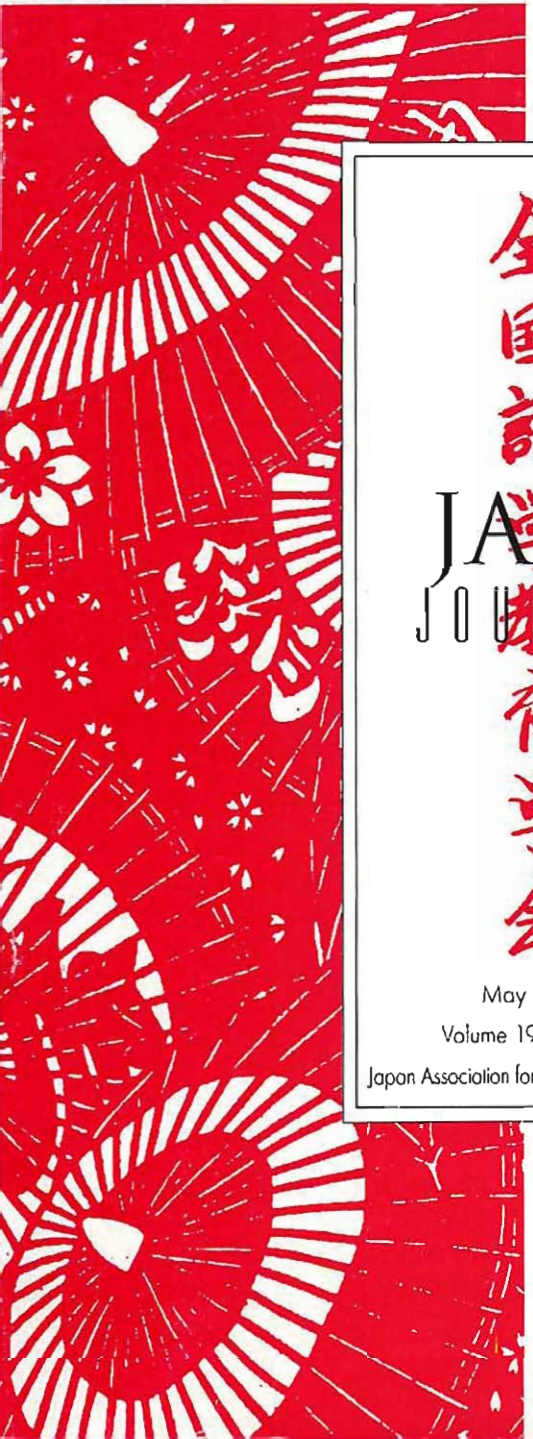
ISSN 0287-2420 ¥950



全国語学教育学会

JALT JOURNAL

May 1997
Volume 19 • No. 1
Japan Association for Language Teaching

Inside this issue:

Communicative strategies • *Eiken* Test
Critical thinking • Letter writing
Performance activities
Test-type variability • Codeswitching
Television Commercials

Contents

May 1997
Volume 19 • No. 1

全国語学教育学会
JALT JOURNAL

Japan Association for Language Teaching

JALT is a professional organization dedicated to the improvement of language learning and teaching in Japan, a vehicle for the exchange of new ideas and techniques, and a means of keeping abreast of new developments in a rapidly changing field. Formed in 1976, JALT has an international membership of more than 4000. There are currently 38 JALT chapters throughout Japan. It is the Japan affiliate of International TESOL (Teachers of English to Speakers of Other Languages) and a branch of IATEFL (International Association of Teachers of English as a Foreign Language).

JALT publishes *JALT Journal*, *The Language Teacher* (a monthly magazine of articles and announcements on professional concerns), *JALT Applied Materials* (a monograph series), and JALT International Conference proceedings.

The JALT International Conference on Language Teacher/Learning and Educational Materials Exposition attracts some 2000 participants annually. Local meetings are held by each JALT chapter, and National Special Interest Groups disseminate information on specific concerns. JALT also sponsors specials events, such as conferences on specific themes.

JALT provides awards for Research Grants and Development, announced annually at the conference.

Membership, open to those interested in language education, includes enrollment in the nearest chapter, copies of JALT publications, and reduced admission to JALT sponsored events. For information on membership, contact the JALT Central Office.

JALT National Officers (1997)

President: Gene van Troyer

Vice President: Brendan Lyons

Treasurer: Larry J. Cisar

Recording Secretary: Jim Chambers

Programs: Carolyn Latham

Membership: Richard Marshall

Public Relations: Jane Hoelker

Chapters

Akita, Chiba, East Hokkaido (affiliate), Fukui, Fukuoka, Gunma, Hamamatsu, Himeji, Hiroshima, Hokkaido, Ibaraki, Iwate, Kagawa, Kagoshima, Kanazawa, Kitakyushu (affiliate), Kobe, Kochi, Kumamoto (forming) Kyoto, Matsuyama, Miyazaki (affiliate), Nagano, Nagasaki, Nagoya, Nara, Niigata, Okayama, Okinawa, Omiya, Osaka, Sendai, Shizuoka, Suwa, Tochigi, Tokushima, Tokyo, Toyohashi, West Tokyo, Yamagata, Yamaguchi, Yokohama

National Special Interest Groups

Bilingualism; Computer Assisted Language Learning; College and University Educators; Global Issues in Language Education; Japanese as a Second Language; Junior and Senior High School Teaching; Learner Development; Material Writers; Other Language Educators; Professionalism, Administration and Leadership in Education; Second Language Literacy; Teacher Education; Teaching Children; Testing and Evaluation; Video

JALT Central Office

Junko Fujio, Office Manager
Urban Edge Building, 5th Floor
1-37-9 Taito, Taito-ku, Tokyo 110 Japan

JALT Journal

Volume 19, No. 1
May, 1997

In this issue

Articles

Five articles are included in this issue. **Ryu Kitajima** examines the influence of a meaning-focused versus a form-focused learning context upon three groups of learners' choices of communication strategies and their overall communicative performance. Results indicate that subjects receiving meaning-focused instruction significantly reduced their use of reduction strategies and increased their use of achievement strategies.

Laura MacGregor undertakes an examination of the *Eiken* tests, administered by the *Eigo Kentei Kyokai* (*Eikyo*), or in English the Society for Testing English Proficiency (STEP). Her study seeks to further understand the *Eiken* tests through 1) an examination of *Eikyo's* statements about the tests and 2) administration of one pre-second level test to a group of Japanese college students and use of these results for studies of the test's reliability and validity.

The importance of promoting thinking has received considerable attention, however, research has not established that training in thinking skills can be effectively combined with EFL/ESL instruction. **Bruce W. Davidson** and **Rodney A. Dunham** report on the use of the Ennis-Weir Critical Thinking Essay Test to assess progress in critical thinking after academic instruction. Results indicate that critical thinking skills can be taught in the EFL classroom.

The relationship between Japanese college students' proficiency in English and their cultural awareness toward a target-language culture (American) are examined by **Taeko Kamimura** and **Kyoko Oi**. Results show that subjects with both high English proficiency and cultural awareness manifested the rhetorical patterns closest to those of native speakers, indicating that cultural awareness may be as important an element as English ability in student writing.

One article in Japanese is included in this issue. **Yoshitaka Kato**, examining the outcome of project work by advanced learners of Japanese at a Russian university, concludes from his study that learner interviews of NSs add to and reinforce the learners' vocabulary and understanding of dialects, and help learners become aware of the way Japanese view Russian customs.

Research Forum

Two articles are included. The first, by **Akihiro Ito**, investigates the effects of differences of test-types on the accuracy rates in interlanguage performance of Japanese EFL learners. The second, by **Ethel Ogane**, looks at codeswitching by adult Japanese EFL learners, suggesting that codeswitching with these learners is discourse related.

Perspectives

Paul Stapleton examines the cultural values of Japanese and American television commercials and suggests they can provide a rich source of both cultural information for teachers interested in promoting learner awareness of target culture values.

Reviews

Texts on multinlingualism in Japan, classroom language use, conversational analysis, grammar, and the use of computer-stored corpora, as well as one learner dictionary, are reviewed by **Steve McCarthy**, **Brian C. Damian Lucantonio**, **Hugh Molloy**, **Kevin Varden**, **Randall David**, and **Charles Adamson**.

From the Editors

JALT On-Line

JALT has expanded its presence on the World Wide Web with several web sites. JALT general information can be obtained at <http://langue.hyper.chubu.ac.jp/jalt/>. Information from JALT's monthly magazine, *The Language Teacher*, is available at <http://langue.hyper.chubu. ac.jp/jalt/pub/tlt/>.

Conference News

The **23rd Annual JALT International Conference** on Language Teaching/Learning and Educational Materials Exposition is scheduled for October 9 to 12, 1997 in Hamamatsu. The theme of the 1997 JALT Conference is Trends and Transition. Those interested in attending the conference are urged to contact the JALT Central Office for more information or visit the JALT '97 web page at <http://www.miyazaki-mic.ac.jp/jalt/JALT97.html>.

The **3rd Annual Pacific Second Language Research Forum**, scheduled for March 26 to 29, 1998, invites proposals for papers and reports of work in progress on data-based research into second language acquisition and use. Papers and reports may be given in either English or Japanese. The deadline for submission of proposals is September 1, 1997. Participants are being sought for symposia on: models of second language processing; language attrition; Japanese as a second language; fossilization; task design and interlanguage; focus on form in instructed SLA; bilingualism; child second language acquisition; psycholinguistic perspectives on the relationship between L2 reading and listening; SLA studies in generative grammar, and cognition and SLA. Proposals should be sent to: Peter Robinson, Aoyama Gakuin University, Dept. of English (PacSLRF '98), 4-4-25 Shibuya, Shibuya-ku, Tokyo 150 Japan. Information about PacSLRF '98 or proposal submission requirements can be obtained from the PacSLRF '98 web site <http://www.als.aoyama.ac.jp/pacslrf.html>.

Corrections

Editorial Advisory Board member **Anne Ediger**'s affiliation was incorrectly listed in *JALT Journal, 18*(2). Her correct affiliation is Teachers College, Columbia University.

Additional Reader **Stephen M. Ryan**'s name was incorrectly given in *JALT Journal, 18*(2).

We sincerely appreciate their efforts on JALT's behalf and wish to apologize for any inconvenience this may have caused.

Articles

Influence of Learning Context on Learners' Use of Communication Strategies[1]

Ryu Kitajima
San Diego State University

This study examines the influence of a meaning-focused versus a form-focused learning context upon learners' choices of communication strategies and their overall communicative performance. For this purpose, an 11-week study was conducted with three groups of students at a university in Japan. The control group studied English in a form-focused learning context. The experimental groups (1 and 2) studied English in a meaning-focused learning context. In addition, experimental group 2 received training in communication strategies. Two types of communicative tasks, 1) description of pictures and 2) narration of a picture story, were administered before and after treatments. Communication strategies were identified by two raters, based on the subjects' audio-taped communicative performance and immediate retrospection. Performances were evaluated separately by the same raters. The results show that both experimental groups significantly reduced use of reduction strategies and increased use of achievement strategies. Similarly, results show that communicative performance by the experimental groups was evaluated higher than that of the control group on the post-test. The findings suggest that learning context has an important influence on learners' use of communication strategies and their communicative performance.

本研究は、意味重視の学習と形式重視の学習が学習者のコミュニケーション法略の選択ならびに全般的コミュニケーションのパフォーマンスに与える影響を考察する。日本の大学において3つのグループが11週間にわたって調査された。統制群は形式重視の文脈で英語を学習し、二つの実験群は意味重視の文脈で学習した。実験群の一つはコミュニケーション法略のトレーニングも受けた。2種類のコミュニカティブなタスクによるテストが事前、事後に実施された。タスクの音声録音とタスク遂行直後の被験者の内省をもとに、コミュニケーション法略が二人の判定者により認定された。パフォーマンスの評価は別途行った。その結果、二つの実験群で、縮小法略が有意に減少し、到達法略の使用が増加した。さらに、事後テストにおいて実験群のパフォーマンスは統制群より高く評価された。これらの結果は、学習の文脈がコミュニケーション法略の使用と、コミュニケーションのパフォーマンスに重要な影響を与えることを示している。

During the past two decades, second language (L2) research has focused on learners' strategies for learning and using a language. One such area involves studies of communication strategies (CSs) (Bialystok, 1990; Bongaerts, Kellerman & Bentlage, 1987; Bongaerts & Poulisse, 1989; Chen, 1990; Dornyei, 1995; Faerch & Kasper, 1983b; Paribakht, 1985; Poulisse & Schils, 1989; Tarone, 1983). CSs, which are learners' attempts to solve communicative problems occurring in the middle of realizing a certain meaning in spontaneous communication, are considered indispensable components of communicative performance. Since learners in communicative situations often encounter difficulty in realizing intended meaning because of limited linguistic resources, they resort to CSs in order to maintain communication.

CSs are generally defined either from an interactionist view (Tarone, 1983) or a psycholinguistic view (Faerch & Kasper, 1983a, 1983b). While the interactionist view emphasizes the learner and the interlocutor's mutual attempts to solve a communicative problem through negotiation of meaning, the psycholinguistic view narrows CSs to those within the learner's internal mechanism. It sets up two criteria for its operation: 1) the existence of a problem which the learner faces for actualizing intended meaning in the target language, caused by insufficient linguistic resources or by the difficulty of retrieving relevant linguistic items from the memory system, and 2) the learner's awareness of the existence of the problem and the necessity of solving it (Faerch & Kasper, 1983a). This study adopts the second definition.

CS studies have found that the use of some CSs results in more effective problem solving than others (Corder, 1983; Faerch & Kasper, 1983a; Tarone, 1983) and that the use of CSs contributes to variation in the overall effectiveness of learners' communicative performance (Chen, 1990). For example, reduction strategies such as *Message Abandonment* (Tarone, 1983) direct the learner to avoid solving a problem and to give up on conveying the message. On the other hand, achievement strategies such as *Analytic* and *Holistic* strategies (Poulisse, 1987) direct the learner to work on an alternative plan for reaching the original goal by means of whatever resources are available. These findings suggest that, in order to understand the variability in learners' communicative abilities in a language, we must study the factors which contribute to the differences in learners' uses of CSs.

Cognitive views of L2 learning (Bialystok, 1990; Bialystok & Sharwood-Smith, 1985; Ellis, 1986; O'Malley & Chamot, 1990; Tarone, 1983) hold that the ability to use language does not come about as a result of an increase in the static rules of the language. Rather, it requires cognitive

strategies and processes which control the knowledge. These cognitive views consider the type of language task a major factor in the proceduralization of language knowledge, by requiring particular cognitive strategies for effective performance.

These views raise questions in terms of the role of learners' strategies in mediating a learning context and their learning outcome. Does a learning context which directs learners to engage in similar tasks for an extended period of time effect acquisition of particular strategies? Do learners in a particular learning context search, use, and acquire particular strategies to process information effectively and efficiently?

The Study

This study investigates whether learners exposed to meaning-focused learning contexts use CSs in communicative tasks differently from those in form-focused learning contexts.

Method

Subjects: An 11-week study was conducted with 15 students of English at a four-year national women's university in Japan. Of the 15 students, 13 were sophomores, one a junior, and one a freshman. Majors were home economics (four), humanities (three), biology (two), and architecture (one). The Ss had studied English for more than six years in required junior and senior high school classes. Though the Ss were taking an English course, as required by the university when they participated in this study, the courses available focused on reading translation, not on English for communication. All Ss voluntarily participated in the study to improve their listening and speaking skills. The Ss English proficiency was considered at least intermediate in terms of grammar in that they had achieved scores high enough for admission on the English portion of the standardized entrance exam administered by the Japanese Ministry of Education. However, all Ss demonstrated difficulty in expressing even simple meanings in spoken English on the pre-test (discussed below). Ss were randomly assigned to either the control group or one of the two experimental groups (Ex1 and Ex2), with five Ss in each group.

Treatment: The control group studied English in a form-focused instructional context. The experimental groups studied English in meaning-focused instructional contexts. In Ex1, CSs were allowed to evolve without explicit instruction. To test the hypothesis that CSs may more efficiently develop

when they are explicitly recognized and practiced, Ex2 received direct instruction on CSs in addition to the activities used with Ex1. All groups received the same amount of instruction: five days a week, 90 minutes per day during the 11-week study. The time allotted for explicit CS instruction (Ex2) was approximately 10 minutes per day.

Teaching method for Ex1: Ss in Ex1 studied English using a series of tasks directing them to express meaning in English without anticipation of the linguistic forms they were supposed to use. Tasks were comprised of communicative exercises such as story telling, discussions, debates on different topics, and picture descriptions. (See Appendix for examples of the communicative activities.)

Teaching method for Ex2: In addition to using the same tasks as Ex1, Ss in Ex2 received explicit instruction about CSs. Analytic and holistic strategies are considered effective ways to convey meaning when the exact words to express the meaning are not immediately available (Poulisse, 1987). Therefore, these two types of CSs were selected for strategy training. On the first day of class, the Ss were introduced to definitions and examples of CSs. In addition, at the beginning of each class, Ex2 spent 5 to 7 minutes solving a lexical problem given on the blackboard. When the Ss encountered difficulty in expressing intended meaning, they were encouraged to solve the problems using the CSs they were studying.

Teaching method for the control group: The control group studied English in a form-focused learning context similar to traditional English classes in Japan. Instruction for this group focused on explicit explanation of particular linguistic forms followed by activities and tasks to practice those forms. The materials used with the two experimental groups (see Appendix) were adapted for use with the control group. With the audio-taped materials, different types of activities (multiple-choice, blank-filling, translation, dictation) were prepared. With the video materials, Ss were given the English transcripts and asked to translate them into Japanese. With the reading passages for debate, the major task for the control group was to translate the reading passages. Similarly, the pictures were used to teach formulaic expressions, with Ss tested on these formulaic expressions orally the following day.

Data Collection

Prior to (pre-test) and immediately after the instructional sessions (post-test), Ss were scheduled to individually perform two kinds of communicative tasks. Each session took approximately one hour. The first task was to describe ten concrete objects drawn on ten separate sheets of paper. Each object to be described was presented on a sheet of paper with a group of

additional objects which shared characteristics with the target object and with one another. For example, a target object, a watering can, was presented with objects such as a bucket, a pitcher, a tin can, and a garbage can. Ss were told that native speakers of English would listen to their audio-taped descriptions to identify the objects described. Therefore, they were encouraged to describe the objects as specifically as possible, particularly if they did not know the exact words to name them. The second task was to narrate a story presented as uncaptioned cartoons. While the Ss were performing these tasks, the researcher was present and encouraged them to talk, by nodding or inserting words such as *Oh, I see*, *Say more*, and *Don't give up*. All performances on the two tasks were audio-taped and transcribed for subsequent analysis.

In addition, each subject's immediate retrospection on task performance was collected in Japanese, following the techniques proposed by Faerch and Kasper (1987). Immediately after the completion of a task, the Ss' audio-taped performances were played back and they were asked to describe what they were thinking while performing the task. The immediate retrospections were audio-taped, transcribed, and translated into English.

Based on the audio-taped and transcribed performances and the transcribed and translated immediate retrospection of the Ss, two raters categorized the Ss' use of CSs, according to the taxonomy described below.

Communicative Strategies Examined

The following CSs, based on existing typologies of CSs proposed by various researchers (Faerch & Kasper, 1983a; Poulisse, 1987; Tarone, 1983), were identified and scored in the analysis process of this study. The distinction between reduction strategies and achievement strategies was considered important in observing how closely learners' communicative goals were achieved (Varadi, 1983).

Taxonomy of Css

Reduction Strategies	Achievement Strategies
Message Abandonment	Analytic
Meaning Replacement	Holistic
	Conceptual Transfer
	Morphological Creativity
	Linguistic Transfer

This study adopted Tarone's (1983) definition of message abandonment: "The learner begins to talk about a concept but is unable to continue

and stops in mid-utterance" (p. 63). The following is an example of message abandonment: *When we hold the . . . I'm sorry I can't.* In this case, a subject intended to describe a *pot holder* but then gave up her attempt when she came across another unknown lexical item, *pot.*

Meaning replacement is different from message abandonment, as Faerch and Kasper (1983a) explain:

> [T]he learner, when confronted by a planning or retrieval problem, operates within the intended propositional content and preserves the topic but refers to it by means of a more general expression. (p. 44)

Based on this definition, meaning replacement strategies were identified when a subject did not entirely give up the problem-solving process, but the realized meaning was far from the communicative goal. For example, I *use this . . . use bake hot cake.* The subject attempted to convey the meaning *spatula*, by describing its functional characteristics, a tool to turn over hot cakes on a frying pan. In this process, the subject encountered other lexical problems such as *turn over.* Instead of giving up, the subject overgeneralized the meaning by simply using the word *bake.*

The category of achievement strategies was based on a typology proposed by Poulisse (1987). The first, analytic strategies, describes characteristic features of a referent to be expressed. An example of this strategy would be, *I use it to clean on the desk or bookshelf to get rid of the dust on the desk.* Not knowing the exact word *duster*, the subject tried to describe the object by presenting its functions. The second category, holistic strategies, are defined as tactics for manipulating a concept and referring to it by using the word for a related concept that shares similar features. For example, another student did not know the exact words *spiral shape*, so she substituted, *This is like a spring.* The next sub-category, conceptual transfer strategies, involve the application of an L1-based concept to refer to a concept in the target language. For example, while not knowing the word *sting* for a description of a wasp, one subject applied the Japanese concept of sting, *needle*, saying *this insect has needle and stick us.*

A fourth kind of achievement strategy (Poulisse, 1987) is morphological creativity, in which learners replace a morphological fragment with a creative one when they do not know the exact word representing a reference. For example, when unable to recall the word *pleased*, a student created a new word by adding a morpheme to the noun *pleasure*, saying *he is very pleasured.* Finally, linguistic transfer refers to the strategy of switching from the target language to the learner's first language to cope with a communicative problem.

In addition to the typology described above, CSs were also identified as operating either at the macro- or at the micro-level. When CSs were used to achieve a global communicative goal, they were considered to be operating at the macro-level; when CSs were used to compensate for a particular lexical item in the process of solving a global communicative problem, they were considered micro-level.

On Task 1, the global goal was to tell what the target objects were. CSs used to solve the problems at this level were regarded as being on a macro level, as in:

> *This is a tool to clean up a room. We drop a dust from our furniture from this tool. This tool have a long handle and the top that has a cloth or wing or something.*

In order to clarify the identity of the object *duster*, this subject described its function and appearance; applying macro-level CSs. On the other hand, to deal with the lexical problem *feather* in the description of the appearance of the duster, the subject used the word *wing*. This was regarded as a micro-level strategy.

On Task 2, CSs used to cope with difficulties in presenting a situation or an action were considered to be on a macro-level. In one of the cartoon stories, a man was lying in the space between two cars parked at the side of a street. The subject, when unable to express the meaning *a man was lying down on a space*, omitted this meaning unit. This strategy was considered macro-level. On the other hand, some CSs used to compensate for particular lexical items such as the use of the general term *doctor*, in place of *archeologist* were considered micro-level.

All CSs used by a subject to realize an intended meaning were counted separately. Two raters independently identified and classified CSs on the basis of their common characteristics, following the taxonomy of CSs established for this study. The results of the raters' identification and scoring showed reasonably good agreement. An average of 73 percent inter-rater reliability was obtained. For those instances in which agreement was not reached, a face-to-face meeting of the raters was held to resolve the difference.

Evaluation of Subjects' Communicative Performance

Besides the identification of CSs, Ss' audio-taped communicative performances were evaluated by the two raters. On Task 1, raters were asked to identify the items described by the Ss. The effectiveness of each subjects' performance was evaluated according to the number of correct objects identified by the raters.

On Task 2, the raters were asked to evaluate the subject's audio-taped communicative performances holistically, on a scale ranging from 0 to 6, based on the amount of relevant information provided by the subject. Inter-rater reliabilities on this evaluation averaged 80 percent, which was considered reasonably high.

Results

The discussion of results focuses on discussion of message abandonment, meaning replacement, analytic, and holistic strategies. Among the CSs listed in the taxonomy for this study, a number of strategies—conceptual transfer, morphological creativity, and linguistic transfer—were infrequently used, both on the pre-test and the post-test. Their sporadic use suggested a preference of an individual subject rather than that of a group. Therefore, use of these strategies will not be discussed here.

Tables 1 – 3 display the raw frequency counts of those strategies used by the Ss, along with means and standard deviations for both Task 1 and Task 2.

Table 1: Descriptive Statistics of Combined Reduction Strategies and Combined Analytic and Holistic Strategies for Task 1 and Task 2

	Task 1				Task 2			
	MA + MR (macro)		A + H (micro)		MA + MR (macro)		A + H (micro)	
	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test
Ex1 Frequency	30	7	20	73	29	6	1	24
M	6	1.4	4	14.6	5.8	1.2	0.2	4.8
SD	1	1.14	1	1.6	1.3	0.8	0.4	2.0
n = 5								
Ex2 Frequency	34	3	20	81	31	0	3	48
M	6.8	0.6	4	16.2	6.2	0	0.6	9.6
SD	2.59	0.89	2.65	3.42	1.4	0	0.5	2.5
n = 5								
CG Frequency	27	31	20	27	32	38	3	13
M	5.4	6.2	4	5.4	6.4	7.6	0.6	2.6
SD	2.3	0.84	2.55	1.95	1.5	3.28	0.8	2.07
n = 5								

MA = message abandonment, MR = meaning replacement, A + H = analytic and holistic strategies

Table 2: Descriptive Statistics of Individual Reduction StrAtegies for Task 1 and Task 2

	Task 1				Task 2			
	MA + MR (macro)		A + H (micro)		MA + MR (macro)		A + H (micro)	
	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test
Ex1 Frequency	11	0	19	7	23	1	6	5
M	2.2	0	3.8	1.4	4.6	0.2	1.2	1
SD	0.84	0	0.83	1.14	0.54	0.44	0.83	0.7
$n = 5$								
Ex2 Frequency	18	0	16	3	28	0	3	0
M	3.6	0	3.2	0.6	5.6	0	0.6	0
SD	3.51	0	2.48	0.89	1.34	0	0.54	0
$n = 5$								
CG Frequency	15	11	12	20	24	25	8	13
M	3	2.2	2.4	4	4.8	5	1.6	2.6
SD	1.87	1.3	1.67	1.58	1.30	3.74	1.14	1.51
$n = 5$								

MA = message abandonment, MR = meaning replacement, A + H = analytic and holistic strategies

Table 3: Descriptive Statistics of Individual Analytic and Holistic Strategies

	Task 1				Task 2			
	MA + MR (macro)		A + H (micro)		MA + MR (macro)		A + H (micro)	
	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test
Ex1 Frequency	17	62	3	11	0	7	1	17
M	3.4	12.4	0.6	2.2	0	1.4	0.2	3.4
SD	1.51	1.51	0.89	1.30	0	0.89	0.44	1.5
$n = 5$								
Ex2 Frequency	16	65	4	15	1	12	2	36
M	3.2	13	0.8	3	0.2	2.4	0.4	7.2
SD	2.28	3.80	0.44	1	0.44	0.89	0.54	1.9
$n = 5$								
CG Frequency	15	20	5	7	0	1	3	12
M	3	4	1	1.4	0	0.2	0.6	2.4
SD	2.35	1.58	0.71	0.89	0	0.44	0.89	2.07
$n = 5$								

MA = message abandonment, MR = meaning replacement, A + H = analytic and holistic strategies

Task 1

Statistical results: All three groups' pre-test scores of the combined macro-level message abandonment and meaning replacement strategies, which were used for problems of target concepts, show no statistically significant difference ($F=0.569$; $df=2, 12$; $p>0.5$). On the post-test, in contrast, the two experimental groups showed a substantial drop in the use of these strategies, while the control group stayed within the range of the pre-test scores. A one-way ANOVA shows that these differences are statistically significant ($F=49.14$; $df=2, 12$; $p<0.01$). Furthermore, the additional comparison between the two experimental groups showed no statistically significant difference ($t=1.24$; $df=8$; $p>0.05$). Individual macro-level reduction strategies show the same trend as the combined macro-level reduction strategies (see Table 2).

The data for macro-level analytic strategies complemented those for the reduction strategies. The means ranged from 3.0 to 3.4 on the pre-test, while on the post-test they ranged from 4.0, for the control group, to 12.4 and 13.0 for Ex1 and Ex2 (see Table 3). Although the control group increased slightly from pre- to post-test, there was a three-fold increase for both Ex1 and Ex2. A one-way ANOVA on the post-test shows this difference to be statistically significant ($F=19.67$; $df=2, 12$; $p<0.01$).

Descriptive results: Reduction strategies were the primary strategies used across the three groups on the pre-test on Task 1. Nevertheless, this does not mean that the Ss did not have knowledge of effective CSs. Their immediate retrospection reveals that they went through the problem-solving process by analyzing the characteristics of the objects. However, they often stopped in the middle of these processes, probably because the communicative problems they experienced seldom appeared in isolation but tended to be interlocked with one another. One subject's performance on Task 1 on the pre-test illustrates. *When we hold the . . . I'm sorry I can't.* The subject's retrospection reveals that she intended to describe the function of the target object, *pot holder.* However, as soon as she came upon another problem with a lexical word *pot*, she gave up her attempt.

On the post-test of Task 1, similar message abandonment behavior was observed among the Ss of the control group, as in this example of a subject's attempt at describing a cylinder: *bottom is circle . . .* (silence). In contrast, Ex1 and Ex2 Ss drastically increased the use of analytic or holistic strategies on the post-test to describe the particular object more specifically and accurately, as seen in the following examples:

(Ex1) The shape is . . . like drink can . . . can drink . . . the bottom . . . object . . . the bottom is circle . . . when you see the side . . . the shape is . . . rectangle.

(Ex2) The bottom . . . the bottom and top is circle . . . circle shape . . . and can juice is . . . the same of this shape.

Task 2

Statistical results: The results concerning macro-level message abandonment strategies, which were used with problems of propositional meaning such as an action and a situation, show similar patterns to those obtained on Task 1. Message abandonment strategies, which constitute the predominant strategies on the pre-test across the three groups, shrank significantly on the experimental groups' post-test performance, while the control group remained within the range of the pre-test scores ($F=8.46$; $df=2, 12$; $p<0.01$). In addition, a t-test showed no statistically significant difference between Ex1 and Ex2 ($t=1.00$; $df=8$; $p>0.05$). On the other hand, there were statistically significant differences among the three groups on the macro-level meaning replacement strategies used on the post-test. A one way ANOVA shows that the Ex2, which had strategy training, eliminated the use of this strategy on the post-test compared to the Ex1, which did not have strategy training, as well as the control group ($F=9.24$; $df=2, 12$; $p<0.01$; $t=3.16$; $df=8$; $p<0.05$).

The results of micro-level achievement strategies, which were used to solve lexical problems within a process of solving a problem of propositional meaning such as an action and a situation, reveal a striking increase in their use among the experimental groups on the post-test. Ex2, which received strategy training, used both micro-level analytic and holistic strategies more than three times as often as the control group and twice as often as Ex1, though no difference was observed among the three groups on the pre-test. A one-way ANOVA as well as a t-test showed the differences among all three groups to be statistically significant ($F=12.97$; $df=2, 12$; $p<0.01$; $t=3.31$; $df=8$; $p<0.05$). The results of individual strategies show similar patterns, though raw frequency counts of holistic strategies were found to be higher than analytic strategies on Task 2 (Analytic: $F=9.33$; $df=2, 12$; $p<$v0.01; Holistic: $F=9.34$; $df=2, 12$; $p<0.01$).

Descriptive results: Reduction strategies, especially message abandonment, were the primary strategies applied by all three groups on the pre-test. While the control group's use of these strategies remained at the same

level as in the pre-test, the experimental groups demonstrated strategic change in their communicative performance on the post-test, with significant increase of their use of micro-level analytic and holistic strategies.

The Ss' performance in the story narration task indicates that the success rate of solving lexical problems depends on top-down as well as bottom-up processing. When solving communicative problems, Ss in the meaning-focused learning context tended to analyze meaning units within the context of the particular communicative task in order to choose an appropriate CS. This context-dependent approach enabled Ss of the experimental groups to express their meaning more accurately and effectively than Ss of the control group, who relied primarily on a context-free approach. One part of the cartoon story on Task 2 was as follows.

> An archaeologist discovered an ancient document on which a statue was drawn. Assuming that it was academically valuable, he decided to search for the statue. After he and his followers endured hardship, they managed to reach their destination, where there were ruins. The archaeologist climbed up one monument and found the statue for which he had been searching.

At this point in the story, there appeared several objects (e.g., *statue, ruins, monument*) which shared some characteristics in this context and therefore needed to be described distinctively in order to make the story coherent. However, most Ss in the control group failed to do so. Following is an example of a control group subject's performance:

> *A man who studied . . . who study . . . old monument . . . he found . . . monument . . . in . . . some . . . place. So, he . . . go to find it . . . he gathered a lot of people to . . . find it with him . . . at last they find a monument . . . and he find the monument he studied.*

In this part of the subject's narration, she used the word *monument* to refer to the statue in the story. However, since there were both statues and monuments in the story, her use of the word was confusing. Rather than just a lack of vocabulary, her problem seems to be also a lack of analysis of the various meaning units within the global context. Objects such as an ancient painting, a figure drawn on it, ruins, monument and statue all were present in the story and therefore should have been referred to with distinct, specific vocabulary items. In fact, Ss of the control group generally failed to analyze and reconstruct meaning elements to represent particular objects in the context. As a result, their performance produced inconsistency and confusion in the story-telling task.

On the other hand, Ss in the meaning-focused learning context managed to distinguish the objects by using different vocabulary items and thus maintained coherence and clarity in their narration of the story. For

example, *a figure* was described as *a doll* or *a monkey. Monument* was described as *pyramid, castle,* or *building made of stone.* The following are two segments of an experimental group subject's narration:

(Ex1) *One famous doctor . . . find . . . a paper. This paper has drawn . . . the doll. He think . . . this is a god . . . of . . . Inca people—ancient Inca people believed . . . God. So, he think . . . in Inka, there is a doll like this . . .*

(Ex2) *There is one man and he studied ancient matter very well and . . . a certain time . . . he found . . . very old picture and . . . there . . . the picture like money . . . is drawn, so he thought . . . South America . . . Ahh . . . there is . . . this picture- . . . he thought . . . same . . . same object . . . he thought there must be same object in South . . . America . . .*

These results suggest the importance of an interactive operation between higher-order interpretive skills at the discourse level and lower-order lexical knowledge in the process of solving lexical problems.

Evaluation of Subjects' Communicative Performance

The independent ratings of identified objects on Task 1 by two raters show that both raters, listening to the Ss' audio-taped description of the target objects without knowing to which group they belonged, could identify two to three out of the 10 objects on the pre-test. In contrast, on the post-test, the same raters could identify an average of eight objects for Ex1 and nine objects for Ex2. These results contrast with the number of objects identified by the same raters from the control group's description on the post-test where the average number remained three objects.

Similar results were obtained from the Ss' communicative performance on Task 2. On this task, the Ss' audio-taped performances were holistically evaluated by the same raters on a scale that ranged from 0 to 6, according to the amount of relevant information provided by the Ss. The average scores given by the raters on the pre-test were 2.6, 2.4, and 2.4 for Ex1, Ex2, and the control group, respectively. Though the control group's scores on the post-test stayed within the pre-test range (2.5), those on the experimental groups' performance improved significantly. The raters scored an average 4.3 for Ex1 and 4.5 for Ex2. One-way ANOVAs yield statistically significant differences between the control group and the experimental groups (rater 1: $F=26.74$; $df=2, 12$; $p<0.01$; rater 2: $F=15.96$; $df=2, 12$; $p<0.01$). On the other hand, no statistical significance was obtained in the comparison between Ex1 and Ex2.

The results of the evaluation of the Ss' communicative performance positively correlated with their use of CSs. The experimental groups' macro-level analytic strategies on task 1 correlated with the success

rates in which the objects they described were identified by the two raters. Similarly, the experimental groups' use of micro-level analytic and holistic strategies resulted in perceptible improvement in the evaluation of effectiveness of their communicative performance on Task 2 on the post-test.

Discussion and Conclusions

The purpose of this study was to examine whether different types of learning contexts would contribute to variation in the use of CSs by learners on communicative tasks. Also, in order to examine whether direct strategy training in a meaning-focused learning context is necessary for learners to develop effective CSs, explicit strategy training was provided for Ex2.

Post-test results did not show generally significant effects of this training, except on the Ss' uses of macro-level meaning replacement and micro-level holistic strategies on Task 2. Ss of Ex2 did pay more attention to the meaning to express, but according to the raters' evaluation there was no significant difference between the performance of Ex2 and that of Ex1 on Task 1 and 2. Therefore, the overall effect of direct strategy training upon learners' choice of CSs was found to be modest. The learning context, then, seems to be responsible for the major effect on the experimental Ss' strategy choice and application.

As a whole, reduction strategies were the primary strategies used by the three groups on the pre-test. In contrast, on Task 1 on the post-test, macro-level analytic strategies became the major strategies for Ex1 and Ex2. On Task 2, micro-level holistic strategies were the most commonly used type of strategy by these two groups. The significant increases in the use of these strategies seem to correlate with the increased effectiveness of these groups' communicative performance. On the post-test, more objects described by Ss of the experimental groups were identified by the raters, and the communicative performance of both groups was judged to be more effective than that of the control group.

It is doubtful that these results were due to the development of the experimental groups' grammatical or lexical knowledge of English. The two groups' task performance (a few examples have been provided previously) clearly indicates that these students had no more linguistic knowledge than the Ss in the control group.

These findings provide further evidence, as suggested by cognitive views of L2 acquisition (Bialystok, 1990; Bialystok & Sharwood-Smith, 1985; Ellis, 1986; O'Malley & Chamot, 1990; Tarone, 1983), that L2 learners' language

ability does not consist of linguistic knowledge alone, but develops as a result of interaction between linguistic knowledge and cognitive processes. Furthermore, the process of applying a particular CS does not operate only at the local lexical level but also includes the analysis of a global communicative goal and the establishment of meaning units within the goal. These findings suggest that strategies used to cope with communicative problems are not automatically transferred from the learner's first language but are acquired in the process of using the target language in particular contexts. While learners may possess effective CSs in their first language, this does not guarantee their being able to apply these strategies to problem-solving in target-language communication.

These findings provide some implication for language pedagogy. If the goal of a language program is to develop learners' communicative ability, a learning context which focuses on explicit instruction of linguistic structures alone may not be sufficient for reaching such a goal. In such a context, learners are likely to develop strategies such as analyzing linguistic structures and memorizing bits of linguistic information but may fail to develop strategies of retrieving those linguistic resources from memory by analyzing a global communicative goal and constructing meaning units within the goal. This may be especially true when the target language is being learned in a foreign language environment where the classroom is the primary source of input for the learners. Therefore, to help learners develop their communicative ability of the target language, various instructional procedures need to be considered.

First, the curriculum should be constructed with a clear goal, as seen in task-based language instruction (Nunan, 1988, 1989), allocating sufficient instructional time to tasks directing learners to engage in problem-solving processes to convey their intended meaning. Second, teaching materials should be chosen and developed with a sound theoretical basis, corresponding with the instructional goal. Instructional materials should include visual aids such as pictures, maps, and symbols which are useful for creating communicative tasks. Video materials also initiate communicative language use such as discussions and debates. Finally, the teaching methods to be adopted must go hand in hand with the goals set up by the curriculum. Learners as active participants in the learning process should be placed in the center of learning, allocating sufficient time for letting them engage in communicative tasks in paired work or small group work.

Finally, in terms of future direction for studies on CSs, the necessity and effect of strategy instruction need further investigation. Some researchers (Bialystok, 1990; Bongaerts, Kellerman & Bentlage, 1987;

Bongaerts & Poulisse, 1989) question the necessity of teaching CSs to learners, from a standpoint that L2 learners already possess those CSs in their native languages and therefore are able to transfer them to L2 communication. On the other hand, some studies (Dornyei, 1995; Wildner-Bassett, 1986) reveal a positive effect of strategy training on learners' communicative performance.

This study, with the inclusion of a strategy training group, examined whether or not a learning context alone is a necessary and sufficient condition for learners to develop effective CSs. The obtained results show no difference between a group with strategy training and the one without it, except on the use of micro-level analytic and holistic strategies on the narration task. However, this was a preliminary study with small sample size and with a short period of time for strategy training. More studies with larger sample size and with more extended period of strategy training need to be conducted before reaching any conclusion. Furthermore, such studies need to incorporate variables influencing the outcome such as learners' proficiency level, learners' personality traits, qualitative aspects of fluency, and different types of discourse in which learners engage.

Ryu Kitajima, Ph.D. Second Language Education, teaches Japanese as a foreign language at San Diego State University.

Note

1. This paper is based on part of the author's doctoral dissertation, which was accepted by SUNY at Buffalo in 1993. An earlier version of this paper was presented at the 28th TESOL convention in Baltimore, March 8-12, 1994.

References

Bialystok, E., & Sharwood-Smith, M. (1985). Interlanguage is not a state of mind: An evaluation of the construct for second language acquisition. *Applied Linguistics, 6*, 101-117.

Bialystok, E. (1990). *Communication strategies*. Cambridge: Basil Blackwell.

Bongaerts, T., Kellerman, E., & Bentlage, A. (1987). Perspective and proficiency in L2 referential communication. *Studies in Second Language Acquisition, 9*, 171-200.

Bongaerts, T., & Poulisse, N. (1989). Communication strategies in L1 and L2: Same or different? *Applied Linguistics, 10*, 253-268.

Chen, S. (1990). A study of communication strategies in interlanguage production by Chinese EFL learners. *Language Learning, 40*(2), 155-187.

Corder, S. (1983). Strategies of communication. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 15-19). London: Longman.

Dornyei, Z. (1995). On the teachability of communication strategies. *TESOL Quarterly, 29*, 55-85.

Ellis, R. 1986. *Understanding second language acquisition*. Oxford: Oxford University Press.

Faerch, C., & Kasper, G. (1983a). Plans and strategies in foreign language communication. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 20-60). London: Longman.

Faerch, C., & Kasper, G. (1983b). On identifying communication strategies in interlanguage production. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 210-238). London: Longman.

Faerch, C., & Kasper, G. (1987). From product to process: Introspective methods in second language research. In C. Faerch & G. Kasper (Eds.), *Introspection in second language research* (pp. 5-23). Clevedon: Multilingual Matters.

Nunan, D. (1988). *Syllabus design*. Oxford: Oxford University Press.

Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge: Cambridge University Press.

O'Malley, M., & Chamot, A. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.

Paribakht, T. (1985). Strategic competence and language proficiency. *Applied Linguistics, 6*, 132-146.

Poulisse, N. (1987). Problems and solutions in the classification of compensatory strategies. *Second Language Research, 3*, 141-153.

Poulisse, N., & Schils, E. (1989). The influences of task- and proficiency-related factors on the use of compensatory strategies: a quantitative analysis. *Language Learning, 39*, 15-48.

Tarone, E. (1983). Strategies of target language learner communication: Message adjustment. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 61-74). London: Longman.

Varadi, T. (1983). Strategies of target language learner communication: Message adjustment. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 79-99). London: Longman.

Wildner-Bassett, M. (1986). Teaching and learning 'polite noises': Improving pragmatic aspects of advanced adult learners' interlanguage. In G. Kasper (Ed.), *Learning, teaching and communication in the foreign language classroom* (pp. 163-178). Aarhus: Aarhus University Press.

(Received April 10, 1996; revised July 25, 1996)

The *Eiken* Test: An Investigation

Laura MacGregor
Sapporo International University

The *Eiken* tests, first administered in 1963 by the Society for Testing English Proficiency (STEP), are highly respected in social, educational, and employment circles and taken by millions each year. However, upon closer scrutiny, it appears that the *Eigo Kentei Kyokai* (*Eikyo*) operates on its own terms. Unlike TOEFL and TOEIC, *Eikyo* does not make information about the tests' reliability or validity available to the public. Therefore, important questions remain unanswered: Are the *Eiken* tests reliable and valid instrument? Do the *Eiken* tests really function as tests of English proficiency? This paper examines the *Eiken* pre-second level test from June 1994. The test was administered to 168 first-year Japanese college students. The results provided data for reliability and validity studies, in an effort to shed light on the value of the *Eiken* pre-second level test as a reputable test instrument of English proficiency. The results of the studies conducted here are far from encouraging with regards to the *Eiken* pre-second level test's reliability and validity.

英検は1963年から英語検定協会によって実施されており、社会的、教育的に、また企業によっても高く評価され、毎年数千人の受験者がいる。しかしながら詳しく調査してみると、英検は独自の方式で実施されており、TOEFLやTOEICと違って信頼性や妥当性についての情報が公開されていない。従って、以下のような重要な疑問は答えのないままである。英検は信頼性と妥当性のあるテストであるのか。英検は本当に英語能力のテストとして機能しているのか。本研究は、英語能力を測定するテストとしての英検の価値を知るために、1994年6月に実施された英検2級のテストを168人の日本人大学一年生に実施し、その信頼性と妥当性を調査した。その結果、英検の信頼性と妥当性ははなはだ心細いものであることがわかった。

Japan is a country whose people thrive on tests, from *kendo* to calligraphy, flower arranging to gift wrapping; tests which evaluate almost every skill imaginable are available for the taking, so to speak. Walk into any bookstore or culture center and you'll see an array of posters and pamphlets advertising such tests. In the academic world, tests are in abundance as well. Entrance examinations which determine students' future high school and post-secondary careers are a fact of life for virtually every family.

By far, the oldest and best established English language tests in Japan are the *Eiken* tests (*Eigo Kentei*), produced by *Nihon Eigo Kentei Kyokai* (*Eikyo*), or in English, STEP: the Society for Testing English Proficiency. Since it began offering the *Eiken* tests in 1963, *Eikyo* has enjoyed a long period of unprecedented success: The Ministry of Education endorses the *Eiken* tests,[1] and recommends students take them. Some schools even offer courses dedicated to *Eiken* test preparation. In the working world, it has made its mark as well. Many employers regard *Eiken* test qualification as a valuable asset and look for it on prospective employees' resumes (MacGregor, 1995). However, despite the test's wide acceptance and use, one important question has been overlooked: Are the Eiken tests reliable and valid instruments to measure English proficiency? Eikyo has failed to give a direct answer. Other language tests, such as TOEFL and TOEIC, publish regular reports with statistical analyses of reliability and validity (TOEFL, 1995; TOEIC, 1995; Woodford, 1980, 1992). Why doesn't *Eikyo*? This question has been raised by other language educators concerned about Eikyo's position as well (Bostwick, 1995; Brown, 1995; Gorsuch, 1995).

This paper seeks answers to the above questions of reliability and validity by conducting analyses on a set of data collected by the author. Because certain reliability analyses are best suited to certain types of tests, it is necessary to begin with the question, is the *Eiken* a criterion-referenced test (CRT) or a norm-referenced test (NRT)? To help answer this question, various aspects of the *Eiken* test will be compared with those of TOEFL and TOEIC. This discussion is followed by an investigation of the *Eiken* pre-second level's[2] reliability and validity, which focuses on four questions: 1) Is the *Eiken* test appropriate for the group *Eikyo* claims to evaluate? 2) Do the test items reflect the practical English found in daily life that *Eikyo* claims to test? 3) Does the test measure the abilities that *Eikyo* claims it does? and 4) Are there any poorly constructed test items? An examination of how the *Eiken* test is scored and how the scores are reported follows. Finally, recommendations are made as to how the *Eiken* test can better serve students and teachers.

Background

The Society for Testing English Proficiency (STEP) was established over 30 years ago as part of a plan by Japan's Ministry of Education to develop education across the nation. Specifically, the goals of the STEP were to popularize and improve the level of practical English in Japan (*Nihon Eigo Kentei Kyokai*, 1994a). In 1963, the same year that the TOEFL

was first administered, the *Eiken* tests were given for the first time to about 37,000 people at three levels (first, second, and third). Five years later, the *Eiken* tests received official approval from the Ministry of Education, which began actively promoting the *Eiken* tests as important tests of English proficiency. With this stamp of approval, the number of test-takers soared, and continues to increase each year.[3] The original three-level test has grown to seven levels: first, pre-first, second, pre-second, third, fourth, and fifth, the first level being the most difficult. The most recent addition to this series was the pre-second level in 1994, introduced to bridge what was felt to be a wide gap in difficulty between the third and second levels.

For levels one to three, the tests are given in two stages. The first stage is a written test (reading and listening comprehension) and the second stage is a speaking (interview) test. Both stages are offered twice a year, in June and October.[4] The focus of this paper is on the first stage of the pre-second level *Eiken* test of June, 1994 (Nihon Eigo Kentei Kyokai, 1994b).

What kind of tests are the Eiken tests?

In the literature on language testing, two types of tests are found (Brown & Yamashita, 1995; Henning, 1987; Hughes, 1989), criterion-referenced tests (CRTs) and norm-referenced tests (NRTs). The differences between the two types lie, not in the actual items themselves, but in the purpose of the tests, how the tests are scored, and how the test scores are used. Therefore, just looking at the test instrument is not enough to determine what kind of test it is.

The purpose of a criterion-referenced test is to evaluate how well the test-taker can perform a specific set of tasks. For example, classroom and term tests evaluate how well a student has learned a defined set of material over a specific period of time. Brown (1996) explained that, "the interpretation of scores on a CRT is considered absolute in the sense that each student's score is meaningful without reference to other students' scores" (p. 2). Therefore, CRT scores do not necessarily conform to a normal distribution.

A norm-referenced test, on the other hand, measures general language abilities. Each student's score is interpreted relative to the scores of all the other students who took the test and the scores generally fall along a normal distribution curve. The TOEFL and TOEIC are norm-referenced tests. The purpose of the TOEFL is to evaluate the English proficiency of foreign nonnative speakers of English, primarily those who intend to study at colleges and universities in the United States or

Canada. Thus, the content of the TOEFL focuses on English for academic purposes. The TOEIC (Test of English for International Communication) is an English language proficiency test which measures how well non-native speakers of English can communicate in English with others in business, commerce, and industry.

Like TOEFL and TOEIC, the *Eiken* tests are English proficiency tests for non-native speakers of English. However, the *Eiken* tests are different in at least two ways. First, the *Eiken* tests are not just one test, but seven different tests. These divisions allow the test-makers to clearly define the material covered, a characteristic of a CRT. *Eikyo* described the contents of its pre-second level test, as follows:

> Successful examinees are able to understand and use general English needed in daily conversation. (High school level; appropriate for a wide range of ages, from high school students to adults in Japan.)
>
> The successful examinee is:
>
> (1) Able to converse about common daily topics. (Able to conduct simple business by telephone; to make easy explanations, leave messages, do simple interpretation, etc.)
>
> (2) Able to read material about common everyday topics. (Able to read news articles, letters, simple pamphlets, etc.)
>
> (3) Able to write about common everyday topics. (Able to write simple letters, notes, memos, etc.) (*Nihon Eigo Kentei Kyokai*, 1994a, p. 8)

Dividing the tests into seven levels is a practical way of handling the wide population, from junior high school to post-graduate levels, that *Eikyo* tests. Further, if students take and pass a test at a level appropriate to their ability, that success is seen to be motivational for their continued study.

The second point in which *Eiken* tests differ from TOEFL and TOEIC is its method of reporting scores. *Eiken* uses a pass/fail reporting system while TOEFL and TOEIC use a converted scale-score reporting system, which makes these two tests "'user-defined' in that scores can be considered in a variety of ways depending upon the requirements of a particular individual or client" (Wilson, 1993, p. 2). Although the reporting styles differ, all these tests follow NRT procedures by using some form of statistical analysis to translate raw scores into standard scores. In other words, none of the tests report their scores as absolute scores.

The question, "What kind of tests are the *Eiken* tests?" remains unanswered. According to Bostwick (1995), "the *Eiken* STEP claims to be a criterion-referenced test in that it specifies proficiency standards and attempts to identify whether the student can pass the pre-established standard" (p. 58). The fact that the *Eiken* tests are divided into seven levels, the

purpose of each level clearly defined by a set of specific tasks, with the language skills required to pass each level specifically defined, gives it qualities of a CRT. However, the way the test is scored (i.e., by converting raw scores to standard scores) is characteristic of an NRT. Therefore, it makes sense to call the *Eiken* tests hybrid CRT/NRT. Knowing that the tests are scored as NRTs is helpful in choosing appropriate reliability measures.

The Study

Method

Materials: The pre-second level test was originally developed for 2nd and 3rd-year high school students (16 and 17-year-olds) (*Nihon Eigo Kentei Kyokai*, 1994c). The most recent statistics show that this group forms the majority of test-takers. For the June, 1996 test, 75% of those who took the pre-second level (227,666 out of a total of 303,955) were senior high school students, 4% (12,471) were junior college students, and 3% (8,549) were university students (*Nihon Eigo Kentei Kyokai*, September, 1996).

There are 75 multiple-choice items on the pre-second level written test. Part 1, which tests vocabulary, idioms, grammar, usage, and reading composition, has 55 items, and Part 2, the listening section, has 20 items. Each item is worth one point for a total of 75 points.

Subjects: The subjects for this study were 182 first-year students (ages 18-20) in five classes at a junior college. Although the reports by *Eikyo* indicate that the pre-second level test is ideally suited to high school students, this higher age level was selected as it best matched the general ability and experience of the group, as outlined below.

A survey, in Japanese, accompanying the *Eiken* pre-second level test to determine the students' experience in taking it, showed that 17% of the students had tried the pre-second level test at least once before but failed, while 40% had never taken an *Eiken* test before. However, because the format of the *Eiken* pre-second level test is similar to high school English tests, it was concluded that lack of *Eiken* experience would not adversely affect the data. Fully 43% had previously passed the third level, confirming that the pre-second level was the most appropriate for this group.

Procedure: In May, 1996, the pre-second level test of June, 1994 was administered to all 182 Ss, along with the survey of *Eiken* experience. After a review of the results of the survey, the results for 14 Ss who had

previously passed pre-second level were eliminated from consideration. The remaining papers (N=168) were scored by hand. The test had 75 items, each worth one point.

Analysis

Analyses to evaluate test reliability were done on the data collected in three categories: Descriptive statistics, item statistics, and consistency estimates. Analyses to evaluate validity were done by comparing the contents of the test with the aims set out by *Eikyo* and the course of study for high school English education. The construction of the individual test items was also examined.

Reliability

Descriptive Statistics: Minimum score, maximum score, midpoint, mean, and standard deviation were calculated. The midpoint is the score which is halfway between the highest and the lowest score. The midpoint, together with the mean, are two statistics which help locate the middle or typical score (Brown, 1996, p. 109).

Item Statistics: Investigation into the reliability of the *Eiken* test began with a look at two types of item statistics: item facility and item discrimination. Item facility (IF) is "a statistical index used to examine the precentage of students who correctly answer a given item" (Brown, 1996, p. 64). The following formula for item facility was used to evaluate individual test items:

$$\text{IF} = \frac{N_1 \text{ (number of examinees who answered correctly)}}{N_2 \text{ (number of examinees who took the test)}}$$

Item discrimination (ID) "indicates the degree to which an item separates the students who performed well from those who performed badly" (Brown, 1996, p. 66). In order to calculate the ID index, it is necessary to differentiate the high scorers from the low scorers. In this study, the upper and lower thirds (33%, or 56 students each) were taken to represent the high scorers and the low scorers respectively. The ID was calculated as follows:

ID = IF (item facility of high scorers) - IF (item facility of low scorers)

That IF and ID are closely connected is apparent from the above formula: If the item is easy (i.e. has a high IF), there should be little discrimination (low ID), and if the item is rather difficult, the discrimination should be high. If the item is too difficult, there should be no

discrimination. According to Brown (1996, p. 69), ideal items in an NRT project have an average IF of .50 and the highest available ID. However, in reality, "items rarely have an IF of exactly .50, so that those that fall in a range between .30 and .70 are usually considered acceptable" (Brown, 1996, pp. 69-70). The items outside this range should either be set aside for revision or discarded. Examination of the ID of the remaining items further evaluates their suitability. The following guidelines were used to evaluate item discrimination:

.40 and up -	very good item
.30 - .39 -	reasonably good, but possibly subject to improvement
.20 - .29 -	marginal item, usually needing and being subject to improvement
below .19 -	poor item, to be rejected or improved by revision (Ebel, 1979, p. 267, cited in Brown, 1996, p. 70)

Consistency Estimates

In this study, the Kuder-Richardson Formula 20 (KR-20) reliability estimate and the standard error of measurement (SEM) were used (Brown, 1996, p. 199, p. 207) to estimate the *Eiken* pre-second level test's reliability. KR-20 was chosen over two other methods of calculating reliability, KR-21 and Cronbach's alpha, for two reasons: 1) it is reported to be the most accurate of the three (Brown, 1996, p. 199), and 2) the results could be compared with the reliability statistics of the TOEFL and TOEIC, which also use KR-20. Further, the *Eiken* pre-second level test follows criteria required by KR-20, in that each item is worth one point and is scored as correct/incorrect.

Test reliability is important because it measures the consistency of the test instrument. If a student takes a test on one day, and then takes the same test again a week or two later, it should produce nearly identical results, that is, if the test is a reliable instrument and little or no learning has taken place between the first and second testing. A test with a reliability coefficient of 1.0 would give precisely the same results for a particular group of test-takers regardless of when it was administered: it would be 100% reliable. Therefore, it is the goal of test-makers to attain the highest possible reliability coefficient.

The standard error of measurement (SEM) is another useful statistic for estimating reliability of NRTs. According to Brown (1996, p. 206), the SEM is "used to determine a band around a student's score within which that student's score would probably fall if the test were administered to him or her repeatedly." A multitude of factors affect test perfor-

mance, only a few of which are connected with the test instrument itself, including items not matched to the purpose of the test, formats unfamiliar to the test-taker, and poorly constructed test items. The majority of factors affecting test performance are related to physical setting (i.e. the test room), and the mental and physical condition of the test taker. With all of these potential distracters, it is sensible to factor in the SEM when determining the cutoff scores.

Validity

A test is a valid instrument if it measures accurately what it claims to measure. For example, an arithmetic test of addition should contain only test items which ask students to add numbers. Further, the test items should not contain ambiguities or misleading information. In the case of the addition test, we can tell whether it is a valid instrument or not just by looking at it. Assessing the validity of English proficiency tests like the *Eiken* is a bit more complicated.

Validity analyses are done on a regular basis by TOEIC and TOEFL. According to a TOEIC report:

> The first validity studies involved the administration of TOEIC to a representative population of Japanese managers, technicians, bankers, and other employees who require English in their work. Researchers compared the candidates' performance on TOEIC to their performance on direct measures of listening, speaking, reading, and writing and determined the correlations. (TOEIC, 1995, p. 3)

Therefore, TOEIC compares its test results to direct four-skills test results to determine test validity. TOEFL conducts a similar type of validity analysis: "TOEFL validation is based upon correlations between test performance of foreign students studying in U.S. colleges and universities and their performance in degree-granting educational programs" (TOEIC, 1995, p. 3). Both TOEFL and TOEIC test for what Woodford calls "concurrent" validity:

> If a language test is supposed to measure whether a person can read Japanese or not then the person who scores high on the test should be able to pick up the Japanese newspaper and tell us what the lead article says. The low scorer should not be able to do it. (1980, p. 4)

It is not known whether *Eikyo* is also doing such validity studies.

Other validation studies involve comparing the results of one test with the results of another. In 1993, TOEIC (Wilson, 1993) published a report of a research study which linked the TOEIC listening section

scores to the scores of the Language Proficiency Interview, a direct assessment of oral language proficiency developed by the Foreign Service Institute of the U.S. Department of State. The numerical correlations between LPI and TOEIC listening sections (.83) proved to be consistently high, suggesting that both tests are, as they claim, effective measures of the ability to understand and use spoken English.

Content validity compares the test specifications with the test contents. If the individual test items match the specifications, then the test can be said to have content validity. This is a subjective evaluation which should be done by a group of testing experts. For the present study, the resources necessary to do the type of validity research described above were not available. Instead, four general questions pertaining to test validity were posed: 1) Is the pre-second level test really appropriate for the group *Eikyo* aims to examine? 2) Do the contents of the test items reflect aspects of "daily life" in Japan, as *Eikyo* claims (*Nihon Eigo Kentei Kyokai*, 1994a)? 3) Do the test items really measure the abilities that they purport to? and 4) Are there any poorly constructed test items?

Results and Discussion

Reliability

Descriptive Statistics: Table 1 shows the descriptive statistics. The midpoint and mean indicate that the typical scores for those who took the test were just above 50%.

Table 1: Descriptive Statistics of the Pre-Second Level *Eiken* Test (N=168; k=75)

Min.	Max.	Midpoint	*M*	*SD*
17	64	40.5 (54%)	38.33 (51%)	9.21

The standard deviation of the test scores was 9.21. Figure 1 shows that the scores of 80 Ss (48%) fell below the mean (0-37 points), and 78 Ss (46%) fell above the mean (39-64 points), while the scores of 10 Ss (6%) were exactly on the mean. The results show that the test performed like a true NRT, conforming to a normal distribution.

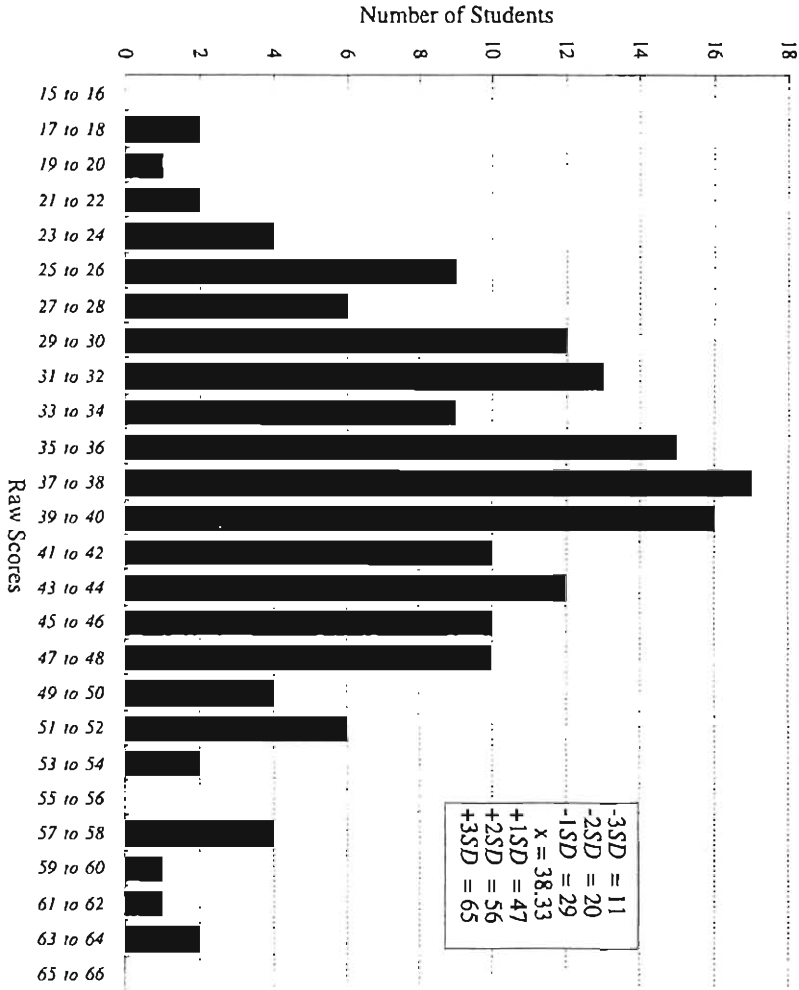


Figure 1: Distribution of Scores on Eiken, Pre-second Level (N=168, x=38.33, *SD* =9.21)

Item Statistics: You will recall that IF shows the percentage of students who answer a given item correctly. For example, 37 Ss answered the first item correctly. Therefore, the IF of item #1 (37/168) is .22 (see Table 2). This means that it was a difficult item because only 22% of the students got it right. Item #3, on the other hand, with an IF of .86, was an easy item for this group.

Table 2: Item Facility (IF) and Item Discrimination (ID) for *Eiken* Pre-Second Grade (N=168)

Item No.	*IF*	*ID*	*Eikyo's IF*†	Item No.	*IF*	*ID*	*Eikyo's IF*†
1	.22	.35	C	41	.81	.27	A
2 •	.36	.41	C	42	.31	.17	C
3	.86	.17	A	43	.72	.44	A
4	.38	.14	C	44•	.43	.30	B
5	.37	.22	D	45	.82	.27	A
6 •	.51	.31	C	46•	.51	.40	B
7 •	.42	.50	C	47•	.60	.37	B
8	.26	.08	D	48•	.55	.44	B
9 •	.50	.42	C	49•	.48	.38	B
10	.23	.14	D	50	.28	.21	C
11	.31	.23	C	51•	.42	.30	B
12•	.57	.37	B	52	.38	.22	C
13	.30	.27	D	53	.39	.16	B
14•	.48	.44	B	54	.18	.13	C
15•	.58	.42	C	55	.28	.16	C
16•	.48	.30	C	56	.60	.17	B
17•	.35	.30	C	57•	.66	.31	B
18	.60	.06	B	58	.89	.21	A
19	.12	.17	D	59	.77	.18	A
20•	.70	.42	B	60	.86	.09	A
21•	.37	.40	C	61•	.57	.37	B
22	.52	.26	B	62	.39	.02	C
23•	.48	.31	B	63•	.49	.40	B
24•	.41	.48	B	64	.65	.16	B
25	.26	.13	D	65	.54	.09	C
26	.80	.07	A	66	.79	.17	A
27	.35	.23	C	67•	.43	.30	B
28	.80	.25	A	68	.63	.37	A
29	.24	.14	D	69	.23	.07	D
30	.52	.21	B	70	.81	.21	A
31•	.55	.51	B	71	.51	.30	C
32	.40	.20	C	72	.94	.03	A
33•	.59	.33	B	73	.71	.19	B
34•	.48	.42	B	74	.26	.21	C
35	.86	.25	A	75•	.54	.45	B
36•	.68	.30	B				
37	.72	.23	A				
38•	.55	.35	B				
39	.48	.28	C				
40	.25	.25	D				

•= Good test items

† Note: IF reported in *Nihon Eigo Kentei Kyokai*, July, 1996, p. 30.

Item #2 (Table 2), has an IF of .36, and an ID of .41, and therefore meets the above IF and ID criteria. It is fairly well centered and discriminates well between high- and low-scoring students. Other items in the test which could be called good test items are indicated by an asterisk. However, there are many items which discriminate poorly. In fact, more than half of the items (59%) do not meet the ID and IF requirements of good test items, falling outside the acceptable range of .30 to .70. Based on these results, further refinements and improvements of many of the test items are needed.

Eikyo published a table approximating IF by assigning letter values to five ranges (Table 2): A=.8–1.0; B=.6–.79; C=.4–.59; D=.2–.39; E=0–.19 (*Nihon Eigo Kentei Kyokai*, July, 1996, p. 30). Because the ranges are so wide (0.19 points each), however, it is difficult to draw any conclusions about IF from these statistics. Besides, IF alone only tells what fraction of the group got the item correct. In order to draw any concrete conclusions about whether the item is functioning well, ID statistics are needed.

Consistency Estimates: The KR-20 result for this study was .82. It is important to keep in mind that this figure of 82% reliability is based on the in-house scores of this sample of 168 Ss. The actual test population, a much greater number with a broader age range, would have had a different reliability index. Unfortunately, this information has not been made available by *Eikyo*. Only a hint as to the reliability of the *Eiken* test was made by an *Eikyo* representative of the test development section, who said that in the years up to 1992, the reliability of *Eiken* tests was between .80 and .90 (*Eikyo* representative, name withheld, personal communication, July 25, 1995). However, he did not divulge the type of analysis done. In any case, this information is not directly relevant to the present study, which deals with the pre-second level test, first introduced in 1994. It is worth noting that in 1989 and 1990, test reliability for TOEIC using the K-R20 formula was .96 (Woodford, 1992).

The SEM for this test administration was 3.9, meaning that the band around which a student's score should be considered is ±4. Therefore, if a subject who scored 37 on the test were to take the test repeatedly, the scores could vary between 33 (-4) and 41 (+4). The passing level for the June 1996 pre-second level test set by *Eiken* was 38 and above, meaning a subject with 37 would have failed. With no apparent margin for error, all it takes is to be one point short to fail.

Validity

Four general questions pertaining to the test validity were posed. First, is the pre-second level test appropriate for the group *Eikyo* claims?

To determine if the test items were suitable for senior high school level students and above, the items of the June, 1994 pre-second level *Eiken* were compared with the nationally approved senior high school course of study (Wada, 1992) and one Ministry of Education approved textbook, *The Crown English Reading* (Hirano, et al., 1996). This revealed that most of the words and idioms on the pre-second level *Eiken* test are taught at some point during the three years of senior high school. [Exceptions are noted as follows: Section 1, item 10 - *vacant*; item 19 - Bill *came all the way from* Florida; item 20 - *take it easy*; Section 2(B), item 9 - *French, Italian, or Thousand Island salad dressing*; Section 3, item 2 - *Are you having some problems there?*; and Section 4(B), item 9 - *farther inland* (*Nihon Eigo Kentei Kyokai*, 1994b).]

The second question regarding test validity asked whether the contents of the test items reflect aspects of "daily life" in Japan, as *Eikyo* claims (*Nihon Eigo Kentei Kyokai*, 1994a)? At least two do not. The following observations are those of the author, not the results of rigorous evaluations by a team of testing experts.

Section 1, item 10:

When the sign on the door of a rest room says "( )," it means someone is using it.

1 OCCUPIED 2 VACANT 3 LIMITED 4 EMERGENCY

(*Nihon Eigo Kentei Kyokai*, 1994c, p. 9)

This item is problematic because restroom doors in Japan seldom have signs indicating whether the stall is vacant or occupied. (The exception is on airplanes, where restroom doors are equipped with such signs). The chance that students are familiar with the context of this item, in English or in Japanese, is remote.

Section 2(B), item 10:

A "Do you have the receipt?"

B "Well, it was a present, but it's too small."

C "What's the problem?"

D "I'd like to exchange this skirt, please."

1 C-B-A-D 2 A-D-C-B 3 D-C-B-A 4 B-C-D-A

(*Nihon Eigo Kentei Kyokai*, 1994c, p. 12)

This item, which asks students to put the four sentences in sequential order (3 is the correct choice), is problematic for three reasons: First, it is culturally inappropriate as Japanese do not customarily exchange items

of clothing that they have purchased, let alone received as a gift. Second, it is illogical. If the person received the skirt as a gift, it is unlikely that she would have the receipt. The third point is not concerned with item validity as much as consistency. In Section 2(B) there are five items, the first four of which follow a question-answer-question-answer sequence. This item, however, has a statement-question-answer-question structure. The fact that it is different from the other items in this section may be a source of confusion for the test takers. If *Eikyo* intended this confusion, it would be of interest to know the reason.

The third question regarding test validity to be asked is if the items on the test really measure the abilities that they claim to. *Eikyo*'s statement about what the successful pre-second level examinee is able to do, converse, read, and write about daily topics, implies that the *Eiken* tests test all four skills. Only listening and reading are actually tested. The TOEIC is similar to the *Eiken* tests in design, as it too tests only listening and reading. The TOEIC differs from the *Eiken* tests in that it measures listening and reading directly, and speaking and writing indirectly. Validity studies have been done to confirm a high correlation between TOEIC results and speaking and writing skills (Woodford, 1980). As *Eikyo* has not published studies to show correlations between its listening and reading tests to speaking and writing abilities, its claims that the successful examinee is "able to converse about daily topics" and is "able to write about common everyday topics" (*Nihon Eigo Kentei Kyokai*, 1994a) cannot be confirmed.

The final question regarding test validity asked if there were any poorly constructed test items. One example is from Section 3 (item 2):

Helen: What did you want to talk to me about? You sounded so mysterious on the telephone.

Sherri: Sorry, but I wanted to tell you this news face to face. I've decided to move.

Helen: But I thought you liked your neighborhood. (2)?

Sherri: No, everything is fine. I just need a change.

2 i) Are you having some problems there? ii) What's the problem?
iii) Isn't everything fine? iv) Why are you moving?

(*Nihon Eigo Kentei Kyokai*, 1994c, p. 13)

The difficulty with this item is in the first of the four possible responses, which also happens to be the correct answer, "Are you having some problems there?" According to Swan (1995), "*some* is most common in affirmative clauses, while *any* is common in questions and negatives" (p. 548). Further, "we use *some* in questions if we expect people to

answer "Yes," or want to encourage them to say "Yes" such as in offers and requests" (Swan, 1995, p. 548) (i.e., *Would you like some more coffee?*). Because choice '(i)' is a question, not an offer or request, it could be argued that it is inappropriate to use *some*.

Two examples of poorly constructed test items occur in the reading passages in Section 4. In the questions following the reading of "Volunteer Guides at Museums" (see Appendix), Item 5 required too much inferencing to make it a viable item:

(5) Professionals at some museums

1 think volunteers should not be paid.

2 feel they know less about museums than volunteers.

3 dislike volunteers because they know more than the professionals.

4 think volunteers cannot do the work of professionals.

(*Nihon Eigo Kentei Kyokai*, 1994c, p. 14)

The correct answer, number 4 (*Nihon Eigo Kentei Kyokai*, 1994c, p. 22), is primarily based on inference, the only clue in the reading passage being, They [professional scholars] feel that amateurs should not do the work of professionals, and that some volunteers act as if they knew everything. The problem is the interpretation of *cannot* in the answer and *should not* in the reading passage.

The second reading passage in Section 4, entitled, "Rainfall in Australia," is problematic in that the text does not correspond to one of the test items (number 7). The pertinent paragraphs and the test item in question are excerpted below:

Most parts of Australia do not receive enough rainfall. In some places there are long periods when it doesn't rain at all. This lack of rainfall is one of the major reasons why such a large country as Australia has such a small population.

Only one-sixth of the continent—a belt of land along the north, east, and south coasts—receives more than 40 inches of rain a year. The rest receives less than 40 inches, and farther inland are somewhat drier areas that receive between 10 and 20 inches.

(7) Where in Australia do they get more than forty inches of rain?

1 In the center of the south coast.

2 In narrow areas along the coasts.

3 In the areas which have monsoon climates.

4 In wide inland areas.

(*Nihon Eigo Kentei Kyokai*, 1994c, pp. 15-16)

Eikyo states that the correct answer is number 2, which implies that all four coasts receive rainfall. However, the supporting statement in the text is . . . *a belt of land along the north, east, and south coasts—receives more than 40 inches of rain a year*, specifically states only three coasts.

To summarize, it is clear that while most of the vocabulary found on this form of the *Eiken* test is appropriate for the intended examinees, problems of context and item construction make the validity questionable. Without evidence from *Eikyo*, it is difficult to conclude that the *Eiken* pre-second level test is a valid instrument.

How are the Eiken *tests scored?* According to *Eikyo*, the passing score for the pre-second level is "approximately 65%" (*Nihon Eigo Kentei Kyokai*, 1994a, p. 7). However, test score statistics since its introduction in 1994 show that the passing scores are much lower. In 1994, the passing percentages and scores were 55% (41+); in 1995, 56% (42+); and in 1996 (June), 49% (37+). This information is somewhat misleading as it is not made clear that these are standard, not raw, scores.

Eikyo's score reporting system is made an even greater mystery by the fact that students never actually see their test scores. All they receive is a report which states either "pass" or one of three categories of "fail," A, B, C.[5] An *Eikyo* representative explained: "A-level failure encompasses scores up 10 points below the passing score; B is up to 15 points below A, and C covers the remaining scores down to zero" (name withheld, personal correspondence, July 16, 1996). These "guidelines" conflict with a report of the pre-second level test of June, 1994, in which *Eikyo* stated that the passing score was 41; A was 34-40 (7 points below the passing score); B, 27-33 (7 points below A); and C, 26 or lower (*Nihon Eigo Kentei Kyokai*, 1994d). These ranges were consistent for 1995 and 1996 tests as well. The discrepancy between the explanation and the published scoring brings into question the integrity of the reporting system.

The following information (in translation) about how the *Eiken* test is scored was received from an *Eikyo* representative in the Planning Division:

> The passing score for the pre-second level is set at approximately 65%. However, the difficulty of the test inevitably varies from time to time, which leads to adjustment of the passing scores each time a new test is given. In the past, the adjustment of the scores has been done by a thorough item by item analysis, looking at the difficulty of each item [IF], and by using point biserial coefficiency. Recently, *Eikyo* has begun experimenting with another method of analysis, Item Response Theory (IRT), as a replacement for the above-mentioned item analysis procedure. (name withheld, personal correspondence, July 17, 1996)

Point biserial correlation is a calculation which shows item discrimination by computing the correlation between individual item responses and total test scores. Like the ID analysis, an item with a low point biserial coefficient may be discarded from the scoring, resulting in changes in the passing score each time a different test is given. However, there is no indication of whether items have ever been discarded by *Eikyo*.

Item response theory true score equating, also used by TOEFL and TOEIC (TOEFL, 1995, p. 9), converts raw scores to equivalent scaled scores. Although *Eikyo* claims to evaluate its test results using IRT methods, there are no published reports to substantiate these claims.

One final concern regarding test scoring is whether *Eikyo* sets a cutoff for the number of people who can pass. Reports in *Eikyo*'s monthly newsletter, *STEP News*, between 1991 and 1996, and information in their brochure, *The STEP Test?* suggest that this is a possibility. The percentage of people who passed the second level test has been consistent at 18% (1991-1996) and the pre-second level at 30-33% (1994-1996) for a number of years.

Conclusion

The results of this study indicate that the reliability and validity evaluations of the pre-second level *Eiken* test are not favorable. First, the reliability in this study is only .82. Is a test that is 82% reliable good enough? For the uninformed consumer, maybe; for test-makers, definitely not. The validity checks in this study show that the content of the test matches the intended group of test takers, perhaps the test's greatest strength. However, there are problems of clarity and context in the items themselves which need to be corrected. Finally, the item facility (IF) and item discrimination (ID) results in this study indicate that more than half of the test items should be revised or removed as they discriminate only fairly or poorly.

Eikyo has been operating a successful testing business in Japan for more than 30 years. In all likelihood, this trend will continue. However, published reports of studies by *Eikyo* on item construction, reliability, and validity are urgently needed to help consumers become better informed about the test, and to encourage research that would improve the quality of the test so that someday the *Eiken* tests might approach reliability in the high .90s.

Acknowledgments

Special thanks are given to Mikiya Koarai for his assistance with this paper and translations of the communications with Eikyo, *Yuji Ushiro for help deciphering Japanese statistical terms, and two anonymous* JALT Journal *reviewers for their advice.*

Laura MacGregor is an Associate professor at Sapporo International University (formerly Seishu University). She also teaches at Sapporo International University Attached Kindergarten.

Notes

1. The Ministry of Education endorses a total of 15 proficiency tests. In addition to the *Eiken* tests, three others are *Kohitsu Shosha Kentei* (penmanship), *Mohitsu Shosha Kentei* (calligraphy), and *Katei Ryori Gino Kentei* (cooking).
2. The *Eiken* tests are ranked from highest, *i-kyu* (first level), to lowest, *go-kyu* (fifth level). *Kyu* is translated here as level, rather than grade, as more appropriate for the *Eiken* ranking system.
3. The total number of people who took the *Eiken* test from 1990-1994 were as follows: (1990) 2,624,106; (1991) 2,761,771; (1992) 2,830,496; (1993) 2,895,912; and (1994) 3,374,140.
4. There is no second stage interview test for the fourth or fifth levels. For these levels, however the written tests are offered three times a year, in January, June, and October.
5. This style of score reporting is not unique to the *Eiken* tests. It is also used in the tests of secretarial skills (*Hisho Kentei*), and *kanji* proficiency (*Kanji Noryoku Shiken*).

References

Bostwick, M.R. (1995). Evaluating young EFL learners: Problems and solutions. In J.D. Brown & S.O. Yamashita (Eds.), *Language testing in Japan* (pp. 57-65). Tokyo: The Japan Association for Language Teaching.

Brown, J.D. (1995). Differences between norm-referenced and criterion-referenced tests. In J.D. Brown & S.O. Yamashita (Eds.), *Language testing in Japan* (pp. 12-19). Tokyo: The Japan Association for Language Teaching.

Brown, J.D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall Regents.

Brown, J.D. & Yamashita, S.O. (Eds.). (1995). *Language testing in Japan*. Tokyo: The Japan Association for Language Teaching.

Gorsuch, G. (1995). Tests, testing companies, educators, and students. *The Language Teacher, 19*(10), 37-41.

Henning, G. (1987). *A guide to language testing*. Boston: Heinle & Heinle.

Hirano, K., Homma, N., Tanabe, Y., Ikegami, Y., Matsuka, H., Yamamoto, S., Sugano, D., Schneider, D., & Wilson, B. (1996). *The crown English reading:*

Teacher's manual. Tokyo: Sanseido.

Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.

MacGregor, L. (1995). More on the *Eiken*. *The Language Teacher, 19*(11), 51, 86.

Nihon Eigo Kentei Kyokai. (1994a). *The STEP test?* Tokyo: Author.

Nihon Eigo Kentei Kyokai. (1994b). *Heisei 6-nen dai 1-kai kentei 1-ji shiken jun 2-kyu (1994, Heisei 6, pre-second level test, 1st stage, 1st administration)*. Tokyo: Author.

Nihon Eigo Kentei Kyokai. (1994c). *Jun 2-kyu gaido (Guide to pre-second level)*. Tokyo: Author.

Nihon Eigo Kentei Kyokai. (1994d). *Eiken jun kaijo reporuto (Eiken pre-release report)*. Tokyo: Author.

Nihon Eigo Kentei Kyokai. (1996, July). *STEP News, 367*. Tokyo: Author.

Nihon Eigo Kentei Kyokai. (1996, September). *STEP News, 369*. Tokyo: Author.

Swan, M. (1995). *Practical English usage* (2nd ed.). Oxford: Oxford University Press.

TOEFL. (1995). *TOEFL test & score manual, 1995-96 edition*. Princeton, NJ: Educational Testing Service.

TOEIC. (1995). *Guide for TOEIC users*. Princeton, NJ: Educational Testing Service.

Wada, M. (1992). *Kotogakko gakushu shido yoryo (The course of study for senior high school)*. Tokyo: Kairyudo.

Wilson, K. (1993). Relating TOEIC scores to oral proficiency interview ratings. *TOEIC Research Summaries, 1*. Princeton, NJ: Educational Testing Service.

Woodford, P. (1980, November). *The test for English for international communication*. Paper presented at the English-Speaking Union Conference, London.

Woodford, P. (1992). A historical overview of TOEIC and its mission. *The 35th TOEIC seminar* (pp. 10-15). Tokyo: The Institute for International Business Communication.

(Received Sept. 27, 1996; revised Dec. 17, 1996)

Assessing EFL Student Progress in Critical Thinking With the Ennis-Weir Critical Thinking Essay Test[1]

Bruce W. Davidson
Hokusei Gakuen University

Rodney A. Dunham
Tezukayama College

Recent trends in EFL/ESL have emphasized the importance of promoting thinking as an integral part of English language pedagogy; however, empirical research has not established that training in thinking skills can be effectively combined with EFL/ESL instruction. This study made use of the Ennis-Weir Critical Thinking Essay Test to assess progress in critical thinking after a year of intensive academic English instruction among Japanese students (N=36). A control group received only content-based intensive English instruction, while a treatment group received additional training in critical thinking. The treatment group scored significantly higher on the test (p<.001). The results imply that critical thinking skills can indeed be taught as part of academic EFL/ESL instruction.

近年、第二言語および外国語としての英語教育では思考力の育成を英語教育の一部として扱うことの重要性が強調されている。しかしながら思考技術のトレーニングが英語教育と効果的に統合できるという結果を示す実証的研究はいまだない。本研究では、Ennis-Weir批判的思考作文テストを用いて、36人の日本人学生が一年間の集中的アカデミックな目的のための英語教育を受けた後にどの程度批判的思考力が発達したかを測定した。統制群は内容中心の英語教育のみを受け、実験群はそれに加えて批判的思考法のトレーニングを受けた。実験群の成績は統制群より有意に高かった。この結果は、批判的思考技術がアカデミックな目的のための英語教育の中で教えることができるのを示している。

Since the advent of research into cognitive development, language teachers and linguists generally have recognized the close connection between language learning and thinking processes. In particular, ESL reading research has shown some correlation between ESL reading comprehension and familiarity with the formal or content schemata of English texts (Carrell, 1987). Furthermore, noting the

unreflective character of many language-teaching approaches that only encourage verbal output or passive input, Tarvin and Al-arishi (1991) and Al-arishi (1994) have explored some methods to make language teaching more thoughtful. Similarly, Chamot (1995) has argued from current educational trends promoting higher-order thinking that EFL/ESL teachers also need to turn the classroom into a community of thinkers. Informal observations may indicate that thinking skills can indeed be taught in an EFL/ESL context (Davidson, 1994, 1995). Without formal testing, however, it is difficult to establish this concretely. Though there has been a lot of thought and research devoted to the development of critical thinking skills in native English speaker educational programs, there has been little research in the area of combining critical thinking with EFL/ESL instruction.

Content-based intensive English instruction has also proven to have many advantages and possibilities (Snow & Brinton, 1988). Is one of them the promotion of critical thinking skills through thought-provoking content? It might be expected that such abilities will develop through discussion, reading, or composition about subjects requiring some serious analytical attention; however, Chance's (1986) survey concluded that critical thinking skills do not develop simply as a by-product of the study of specific subjects. In addition, Halpern (1993) cites evidence from various sources that critical thinking skills can be inculcated through explicit instruction.

These issues and findings inspired a pilot study to discover whether or not critical thinking could be taught to Japanese students of English in a content-based EFL program. After defining what we mean by "critical thinking," we will describe the intensive English program the subjects were enrolled in along with the specifics of the current study. Two research questions guided us:

1. On a critical thinking test task, will English learners exposed to critical thinking skills-training do significantly better than similar students who have not received such training?

2. Can a critical thinking test designed for native English speakers be used as an instrument for evaluating critical thinking skills among non-native English learners?

Critical Thinking: Concept and Inventory of Component Skills

Critical thinking involves rational judgment and discernment of the elements of reasoning. Various definitions of critical thinking reflect

this. Norris and Ennis (1989) explain critical thinking as "reasonable and reflective thinking that is focused upon deciding what to believe and do" (p. 3), a definition also stated somewhat differently by Lippman (1991), who defines it as healthy skepticism, and Siegel (1988), who considers the critical thinker to be one who "is appropriately moved by reasons" (p. 2). In contrast to rote memorization or simple information recall, methods for encouraging critical thinking have as their goal the stimulation of the analytical and evaluative processes of the mind (Paul, 1992). Norris and Ennis (1989) have listed a number of critical thinking abilities to develop:

Elementary Clarification
1. Focusing on a question
2. Analyzing arguments
3. Asking and answering questions that clarify and challenge

Basic Support
4. Judging the credibility of a source
5. Making and judging observations

Inference
6. Making and judging deductions
7. Making and judging inductions
8. Making and judging value judgments

Advanced Clarification
9. Defining terms and judging definitions
10. Identifying assumptions

Strategies and Tactics
11. Deciding on an action
12. Interacting with others (p. 14)

This concept of critical thinking and inventory of skills inspired both the instructional treatment and the selection of the Ennis-Weir test in this study.

The Study

Method

Subjects: All participants in the study (N=36) were first-year students enrolled in a private women's junior college in Osaka, Japan. The college's curriculum consisted mainly of an intensive academic English program. Weekly English courses included Oral Discussion (3 hours), Composition (2 hours), Reading (3 hours), Pronunciation (3 hours),

and Grammar/Listening (2 hours), totaling 13 hours a week, considering each 50-minute class session as an hour. Oral Discussion, Reading, and Composition followed a topical syllabus of six units of instruction, which included such themes as Prejudice/ Human Rights, Advertising/ Consumerism, and Women's Issues/Child-raising. The integrated, content-based aspect of the program was meant to involve students in in-depth analysis and expression concerning subjects significant in their own lives and in Japanese society. This course of study would seem well-suited to encouraging the development of critical thinking skills as a by-product, since the topics all necessitate thought. Along with the topic, each unit also introduced a rhetorical mode: Illustration, Process, Definition, Classification, Comparison/Contrast, and Persuasion. The first three composition units required students to write a paragraph using each mode, and the last three progressed to multi-paragraph essays. The persuasion essay was written in a mini-term paper format, with references.

In addition to these integrated intensive English courses, Ss took a weekly one-hour seminar course, also conducted in English and concerning some interesting topic or theme such as *American Holidays* or *Traditional Folk Songs*. The treatment group ($n=17$) was composed of students from a seminar on Critical Thinking.[2] Volunteers not enrolled in the critical thinking seminar ($n=19$) served as a control group.

Ss had varying degrees of English proficiency as measured by an in-house proficiency test. At the beginning of the year, this test divided all students into five levels of classes according to scores: A, B, C, D, and E. The A classes had the highest level of proficiency and included students returning from a year or longer of study abroad, whereas the E classes were much less proficient. Regardless of proficiency, however, all classes received similar instruction based on the same content and rhetorical modes noted previously. These Ss represented a broad range of English proficiency levels, as measured by the in-house test, with both groups containing a similar range. Because of the small number of Ss at each level, with the exception of level C ($n=14$), levels were grouped into A+B, C, and D+E for later comparison. The distribution of Ss among the groups was fairly even (control $n=5$, 10, 4; treatment $n=7$, 4, 6). No pre-test was given, in line with the advice of Ennis and Weir (1985), who state that a pre-test is not necessary in research using the test as long as a control group exists. Babbie (1983) has noted that a post-test-only control group design is quite acceptable as long as group assignment is random. Since students enrolled in seminars through a semi-lottery system, the authors consider that in this case group as-

signment generally embodied the spirit of randomness, although it was not completely random.

Treatment: The treatment group took part in a course designed to train them in basic elements of critical thinking: source credibility, inductive reasoning, informal deductive logic, and assumption-identification. These broad categories encompass most of Norris and Ennis's (1989) list of critical thinking skills, so they were adopted as a framework for the seminar course. During the first semester, instruction dealt with inductive reasoning and source credibility; during the second semester, the emphasis was deductive reasoning and assumption-identification. The course began with an introduction to the concept of critical thinking. Sessions were devoted to exploring various kinds of reasoning fallacies and misuse of evidence, such as over-generalization and the false dilemma (Chaffee, 1990; Damer, 1995). Students were given lists of brief fallacious arguments and asked to explain the problems of each in their own words. In the second half of the semester, the focus shifted to source credibility. Students did exercises in which they evaluated varying accounts of the same event according to differing viewpoints. For example, in groups they discussed and ranked accounts of the results of an international conference, keeping in mind a list of question-criteria: Does the news presenter have a reason to be biased? Is the source an expert in the field? (Beyer, 1991). Students brought in similar examples to present and evaluate.

In the second semester the emphasis shifted to basic argument analysis. First, students did exercises to help them distinguish real arguments from bare claims offering no reason (Engel, 1994). Then they identified the claims and supporting reasons. Later in the semester, the instructor introduced less-obvious aspects of deductive reasoning: unstated assumptions and implications (Scriven, 1976). Using magazine advertisements and other material, students practiced identifying assumptions and implications. As a result of the Hanshin earthquake and other circumstances, a total of only 18 class hours was actually devoted to the course's content.

Instrument: The Ennis-Weir Critical Thinking Essay Test

Test Description: The Ennis-Weir test was chosen for various reasons. One is that it is one of the most generally well-accepted measuring instruments among educators in the critical thinking movement (Walsh & Paul, n.d.), and inter-rater reliabilities have been very high when it has been used (Ennis & Weir, 1985; Hatcher, 1996). Another is that, in contrast to multiple-choice tests, it allows students to justify varying responses, and the test

itself presents a realistic critical evaluation task. Other critical thinking tests are available, but almost all of these are multiple-choice instruments that suffer from various weaknesses such as background bias and the impossibility of knowing the reasoning behind an examinee's answer-choice (Ennis, Millman, & Tomko, 1985; Norris & Ennis, 1989). Furthermore, the relatively simple subject-matter and language of the Ennis-Weir test make it suitable for non-native speakers. It has been used successfully with first-year junior high school native English speakers in the U.S.

The test itself contains a simple set of instructions and a letter to a newspaper editor containing 10 brief paragraphs. The fictional writer, Raywift, recommends that overnight parking be prohibited on all the streets of his town, Moorburg. After a brief introduction, eight numbered paragraphs elaborate the argument. Most are weak and commit various common reasoning fallacies such as equivocation, irrelevancy, poor statistical sampling, and circular reasoning, but some contain legitimate support, consisting in the use of qualified experts or a relevant reason. Point-by-point, the examinee's task is to judge the thinking of each of these numbered paragraphs and to evaluate the strength of the letter's argument as a whole in a final summary-paragraph. For example, in responding to Paragraph 3, examinees are expected to notice that a relevant reason is offered to support Raywift's argument. Similarly, test-takers are supposed to show some indication that they comprehend the flaws of the experiment in Paragraph 6. (See Appendix for copies of the introduction and paragraphs 3, 6, and 8.)

A clear and specific scoring protocol accompanies the test indicating various possible answers and how each is to be scored. Points are awarded both for judging correctly and for indicating a valid reason for

Table 1: Critical Thinking Skills Addressed on the Ennis-Weir Test

Paragraph	Skill
1.	Noticing misuse of analogy and/or shift in meaning
2.	Recognizing irrelevant reasoning
3.*	Recognizing relevant reasoning
4.	Recognizing circularity and/or the lack of a reason
5.	Recognizing defective reasoning
6.	Recognizing insufficient sampling
7.	Recognizing equivocation and/or the use of an arbitrary definition
8.*	Evaluating the credibility of expert testimony

*Paragraphs that exhibit sound reasoning

one's judgment, and a penalty point of -1 can be deducted for poor reasoning. Each answer can receive a maximum of 3 points and a minimum of -1, except for the summary paragraph, where a maximum of 5 points can be awarded. Therefore, the overall score can range from -9 to +29. In general, the protocol gives latitude to raters to award points whenever an examinee can give a credible reason in support of this evaluative judgment, even when it differs from that of the protocol writers. Brief answers are acceptable as long as they indicate a valid judgment, backed up with a sound reason for that judgment.

We considered the possibility that cultural differences between America and Japan might bias the results, but we felt that in this case culture would not be a significant issue. For one thing, the Moorburg letter concerns street parking, which is not a point of significant cultural difference between the U.S. and Japan. Street-parking laws also exist in Japan and are stringently enforced. If anything, parking is more of a problem in Japan, so Japanese readers might be even more likely to identify with a parking-related issue than Americans. Furthermore, writing a letter to an editor complaining about a public problem, as in the Raywift letter, is a common practice in Japan as well. The directness and abrasiveness of Raywift's style are perhaps the only aspects of the letter that might seem strange or unsettling to a Japanese reader. However, it should also be noted that abrasive political rhetoric is not unknown in Japan.

A limited amount of research has been done in the U.S. using the test. The largest study to date has been Hatcher's (1996) at Baker University. Over a period of four years (1990-1994), American freshmen scored an average of 11.8 to 13.8 on the Ennis-Weir test after a year-long compulsory critical thinking course. They had scored from 5.8 to 9.4 on a pretest and registered gains of 2.8, 5.8, 5.8, and 6.0 points. Interestingly, a number of Chinese and Japanese students at Baker University also took part in the study, but their scores were eliminated from it because they consistently scored poorly. Hatcher (1994) speculated that their low scores may be due to Oriental politeness and accordingly a hesitancy to criticize the Moorburg letter.

Test Administration: In the last week of second semester classes, the Ennis-Weir Critical Thinking Essay Test (1985) was administered, with the control group given the test within the same week. Both groups had 80 minutes to read the test and write nine brief paragraphs in response, twice the amount of time recommended by the test-makers. Since the subjects were non-native English speakers, it was felt that

more time would be necessary for them to comprehend the material and compose answers. To help them with the language aspects of the test, they were allowed to use dictionaries. Furthermore, before taking the test, all subjects received two sample test items with model answers to make sure the Ss understood two things: (a) that they had to make a clear evaluative judgment as to whether the argument in each paragraph was a good one or not and (b) that they had to give a clear reason or explanation for their judgment. Without such explicit direction, the subjects might not have done either of these two things. However, students in the present study were used to doing peer-evaluation of essays in composition classes, so the idea of writing comments or criticism about a piece of writing was already somewhat familiar. Basic information about the Ss, including English level and overseas experiences, was collected on the answer form.

Results

Tests were scored blindly and independently by two raters. The test-raters in this study found little difficulty in using the protocol to judge student answers. Grammatical or vocabulary problems were overlooked unless they made an answer incomprehensible. Inter-rater reliability was found to be adequate ($r = .72$). The scores and information collected were examined. Therefore, the average scores given to each student for each of the 10 scores on the test, one score for each of the nine paragraphs and a total score, were used for all subsequent analysis.

The small number of Ss in the study (control $n = 19$ and treatment $n = 17$) makes relationship detection difficult unless it is very strong. Because this is the first study of this type, we were interested in detecting moderate relationships as well as strong ones. Consequently, we decided that the risk of committing a Type I error would be less important than missing moderate relationships. Therefore, the significance level of .10 was chosen as the cutoff for accepting or rejecting relationships. Nevertheless, we have reported here the exact probability for all results that indicated statistically significant relationships.

The most important analysis, of course, dealt with the effect of critical thinking training on test scores. Therefore, we conducted a t test to compare the scores of the two groups. The treatment group's mean score on the Ennis-Weir Test was statistically 6.6, significantly higher than the control group's mean score of 0.6 (t (27.73) = -4.99, $p < .001$). Table 2 shows the range of scores for each group and details the differences. As the table shows, 10 Ss in the treatment group scored 7 or

Table 2: Group Comparative Scores on the Ennis-Weir Test

Score Range	Control Group	Treatment Group
-4.0 to 0.0	9	1
1.0 to 2.0	6	1
2.5 to 6.5	4	5
7.0 to 13.5	0	10
Total *N*	19	17
Mean	0.6	6.6
Median	1.0	7.5
Mode	-1.5	3.0

higher, while only one scored 0 or lower. In contrast, in the control group, nine Ss scored 0 or lower, with no score higher than 6.5.

Next, the individual paragraph scores of the control and treatment groups were compared. There were statistically significant differences between the mean scores of the control group and the treatment group on two paragraphs: the third paragraph (M=-0.18 and 0.50, respectively, t (33.87) = -2.10, p=.043) and the sixth paragraph (M=-0.55 and 0.15, respectively, t (27.10) = -2.59, p = .015). The difference in scores on the eighth paragraph approached statistical significance (t (33.69) = -1.71, p = .096), with the treatment group scoring higher (M=0.47) than the control group (M=-0.52). However, the scores on the remaining paragraphs showed no statistically significant difference between the control and treatment groups (p>.10).

Since the test was in English, a foreign language for the Ss, proficiency may have affected scores. As mentioned, since the number of Ss at each proficiency level, except C (n=14), was quite small (A level n=7, B level n=5, D level n=5, and E level n=5), A level was grouped with B, and D level with E. The C level was left intact, creating a three-level variable. The distribution for this ordinal variable was compared for levels A+B, C, and D+E between the control group (n=5, 10, 4) and the treatment group (n=7, 4, and 6). As determined by a chi square test, the distribution indicates that there was no relationship between English proficiency and the type of group (χ^2(2, N=36) = 3.204, p = .202). An analysis of variance was also run to examine the relationship between English level and test scores. There was no statistically significant relationship between the two variables (F(2, 35) = 1.57, p = .224). Table 3 shows that the range of scores was comparable for each proficiency level. Judging by this analysis as well as the phrasing of student answers on the test, we believe that students gen-

erally did not do poorly simply as a result of an inability to understand the English of the test. The sample test items appeared to succeed in helping students to grasp the kind of test task they were engaged in, and the wording of the test did not appear to present an insurmountable problem even for lower-level Ss.

Table 3: Scores on Ennis-Weir Test by English Level

Score Range	A-B Ss	C Ss	D-E Ss
-4.0 to 0.0	4	4	2
1.0 to 6.5	4	8	4
7.0 to 13.5	4	2	4
Total *N*	12	14	10
Mean	4.3	1.8	4.7
Median	3.8	1.5	4.3
Mode	10.5	1.0	3.0

The next variable examined was overseas experience, since a number of Ss had lived a year or longer in an English-speaking country. Using a *t* test, scores of Ss who had traveled overseas and those who had not were compared. The differences in scores between the two groups was not significant for total scores or for any individual paragraph score ($p > .10$).

Because the Ennis-Weir test deals with parking problems, each of the subjects was asked to report on the test form whether or not she possessed a driver's license. Japanese students often first learn to drive at the age of 19 or 20, the age of the Ss in this study, and familiarity with driving an automobile may have helped some Ss do better on the test, which concerns a parking problem. Scores for the two groups, those with licenses and those without, were compared. Total scores were statistically the same for both groups; however, students without driver's licenses ($M = -0.66$) scored statistically significantly lower on the seventh paragraph than Ss with ($M = -0.17$) [$t\ (21.14) = -1.84$, $p = .079$]. On the eighth paragraph, those without licenses ($M = -0.24$) also scored significantly lower than Ss with ($M = 0.70$) [$t\ (31.50) = -3.31$, $p = .002$]. Otherwise, the fact of having a driver's license showed no significant relationship with student scores ($p > .10$). Since the specific issues addressed in the seventh and eighth paragraphs are not directly related to the experience of driving we consider the statistical significance to be unrelated to the current study.

These statistical analyses appear to indicate that the differences in scores between the treatment and control groups cannot be accounted for by differences in English proficiency levels or other factors such as overseas experience or having a driver's license. Therefore, the differences in scores on the Ennis-Weir Test can probably be attributed to the critical thinking training given the treatment group.

Discussion and Conclusion

Both research questions can be answered in the affirmative, based on the results of this study. The Ennis-Weir test, designed for native English speakers, appears to be usable for non-native English learners. Furthermore, it is encouraging to find that even a small amount of instruction in the basics of critical thinking appeared to result in higher scores for the treatment group. Critical thinking skills can apparently be taught to some extent along with English as a foreign language and can, therefore, enhance a content-based course of study. In view of the relatively small amount of actual instruction, the rather low average score of 6.6 is not surprising, and is much better than the performance of the control group ($M = 0.6$). As a point of comparison, Baker University American freshmen registered gains of 2.8, 5.8, 5.8, and 6.0 (Hatcher, 1996) in four successive years. Interestingly, three of those gains approximate the difference of 6.0 that we found in the mean scores of our two groups, though the mean score of the treatment group (6.6) is only half that of the average post-test scores of the Baker freshmen. Looking at the individual test items, differences between the two groups appeared specifically in items which had received instructional attention in the critical thinking class. Paragraph 6 deals with the misuse of statistics, a reasoning problem dealt with in class, while Paragraph 3 featured a relevant reason, another instructional point. The difference in performance on Paragraph 8, which concerned the use of experts and their credibility as sources, also approached statistical significance, and that area also had received attention in the source-credibility component of the critical thinking seminar. In contrast, little difference in scores appeared in the case of Paragraphs 1 and 7, which both concerned inappropriate definitions, an area not dealt with in the course. Furthermore, there was little difference in scores on Paragraph 4, which consists in circular reasoning and was very similar to one of the sample test items. Perhaps because of its similarity, 35 Ss responded correctly to it.

The overall quality of the answers of the two groups differed, but they shared certain tendencies indicating a general weakness in the area

of critical thinking skills. This is not surprising in view of the fact that Japanese education does not seem to encourage debate or the critical evaluation of reasoning (Davidson, 1995). Detailed consideration of the answers themselves is beyond the scope of this study, but it is revealing to explore the kinds of errors consistently made by the participants. All of the subjects had been taught to identify and use definition, illustration, and argumentation as rhetorical modes; however, this training apparently did not prepare them to recognize reasoning errors related to these modes. For example, in the case of the 24 Ss who positively but incorrectly evaluated Paragraph 2, all gave as their justification the fact that Raywift provided a reason grounded in reality or else that he gave a concrete example. They missed the fact that both the reason and the example were irrelevant to his argument. Similarly, students accepted the definition-arguments in Paragraphs 1 and 7, even though the definitions offered by Raywift were inappropriate. For instance, he argues in Paragraph 7 that his opponents "don't know what 'safe' really means. *Conditions are not safe if there's even the slightest possible chance for an accident*" [italics added] (Ennis & Weir, 1985, p. 13). Only 2 Ss found fault with this impractical definition of the concept of safety; the others credited him with giving a clear definition. Likewise, 25 of 36 Ss accepted the false analogy used in Paragraph 1. Though the treatment group fared better on some paragraphs and in their overall scores, these common tendencies seem to point to a general need for critical thinking training among these particular Japanese EFL students that perhaps is not being addressed adequately by practice in English rhetorical modes or content-based study. It is even possible that exposure to rhetorical modes such as definition, illustration, and argumentation may only predispose students to accepting weak ideas simply because they are presented in the proper rhetorical format. Without concurrent attention to reasoning fallacies and the pitfalls related to each mode, teachers may discover that for their EFL/ESL students, a little bit of knowledge of rhetorical modes is a dangerous thing. Such students may one day find themselves struggling with the reasoning tasks required in an English academic setting, regardless of their general English language proficiency or familiarity with English modes of expression.

Though Hatcher (1994) speculated that politeness and a hesitancy to make negative judgments may have inhibited Japanese and Chinese performance on the Ennis-Weir test at Baker University, answers in our study reveal that they did not err only in positively evaluating weak paragraphs. They also often negatively assessed Raywift's better arguments in Paragraphs 3 and 8. For example, a number of Ss rejected

Raywift's citation of qualified experts in Paragraph 8 as "just opinions." Ss did not appear to suffer from any hesitancy to criticize.

This is only a limited pilot study, and more research of a similar type needs to be done to substantiate these tentative conclusions. Larger student samples are needed. Also, it would be helpful if a translated version of the test could be administered to groups of similar Japanese students to remove completely the possibility that English language deficiencies may to some extent account for the lower scores. For cultural and linguistic reasons, however, such a translated test may be difficult to make and administer. Students in mixed-nationality EFL/ESL programs in other cultural settings could also provide interesting and relevant data about critical thinking abilities and the possibility of developing and testing them in English language programs, since English language-learning problems related to thinking are not confined to Japan. English instructors in other places have noted reasoning weaknesses similar to the ones we have found (Sherman, 1992; Matthews, 1994). Furthermore, it would be informative to experiment with other standard tests of critical thinking in EFL/ESL programs. Finally, it is worth exploring the question of whether training in critical thinking can improve general English language proficiency, especially in writing and reading. Nevertheless, we hope to see the Ennis-Weir test applied by others in studies bearing some similarity to ours. This relatively unexplored area invites further inquiry.

Bruce W. Davidson teaches English at Hokusei Gakuen University in the Department of Social Welfare and researches issues related to critical thinking education. Send correspondence to: z00134@hokusei.ac.jp.

Rodney A. Dunham teaches in the English Department at Tezukayama College and conducts research in mass communication.

Notes

1. Earlier versions of this paper were presented November 5, 1995 at the Twenty-First Annual JALT (Japan Association for Language Teaching) International Conference, Nagoya, and July 29, 1996 at the Sixteenth International Conference on Critical Thinking and Educational Reform, Rohnert Park, CA.
2. The treatment group consisted of 17 out of 22 members from the seminar on critical thinking. Five could not take the Ennis-Weir test due to circumstances arising from the Kobe-Osaka earthquake.

References

Al-arishi, A. (1994). Practical ways to promote reflection in the EFL/ESL classroom. *English Teaching Forum, 32*(2), 2-5.

Babbie, E. (1983). *The practice of social research* (3rd ed.). Belmont, CA: Wadsworth.

Beyer, B. (1991). Practical strategies for the direct teaching of thinking skills. In A. Costa (Ed.), *Developing minds: A resource book for teaching thinking: Vol. 1* (2nd ed., pp. 274-279). Alexandria, VA: Association for Supervision and Curriculum Development.

Carrell, P. (1987). Content and formal schemata in ESL reading. *TESOL Quarterly, 21*, 461-481.

Chaffee, J. (1990). *Thinking critically*. New York: Houghton-Mifflin.

Chamot, A. (1995). Creating a community of thinkers in the ESL/EFL classroom. *TESOL Matters, 5*(5), 1, 16.

Chance (1986). *Thinking in the classroom: A survey of programs*. New York: Teachers College Press, Columbia University.

Damer, E. (1995). *Attacking faulty reasoning: A practical guide to fallacy-free arguments*. Belmont: Wadsworth.

Davidson, B. (1994). Critical thinking: A perspective and prescriptions for language teachers. *The Language Teacher, 18*(4), 20-26.

Davidson, B. (1995). Critical thinking education faces the challenge of Japan. *Inquiry: Critical Thinking Across the Disciplines, 14*(3), 41-53.

Engel, M. (1994). *With good reason: An introduction to informal fallacies*. New York: St. Martin's Press.

Ennis, R., Millman, J., & Tomko, T. (1985). *Cornell Critical Thinking Tests Level X & level Z* [manual]. Pacific Grove, CA: Critical Thinking Press and Software.

Ennis, R., & Weir, E. (1985). *The Ennis-Weir Critical Thinking Essay Test*. Pacific Grove, CA: Critical Thinking Press and Software.

Halpern, D. (1993). Assessing the effectiveness of critical-thinking instruction. *JGE: The Journal of General Education, 42*(4), 238-254.

Hatcher, D. (1994, August). *Grading the Ennis-Weir Critical Thinking Essay Test*. Paper presented at the Fourteenth International Conference on Critical Thinking and Educational Reform, Rohnert Park, CA.

Hatcher, D. (1996). Combining critical thinking and written composition: The whole is greater than the sum of the parts. *Inquiry: Critical Thinking Across the Disciplines, 15*(2), 20-36.

Lippman, M. (1991). *Thinking in education*. Cambridge: Cambridge University Press.

Matthews, C. (1994, March). *Using news stories to develop thinking/reasoning skills*. Paper presented at the TESOL annual convention, Baltimore.

Norris, S. & Ennis, R. (1989). *Evaluating critical thinking*. Pacific Grove, CA: Critical Thinking Press and Software.

Paul, R. W. (1992). *Critical thinking: What every person needs to survive in a rapidly changing world*. Santa Rosa, CA: Foundation for Critical Thinking.

Scriven, M. (1976). *Reasoning*. New York: McGraw-Hill.

Sherman, J. (1992). Your own thoughts in your own words. *ELT Journal, 46*(2), 190-198.

Siegel, H. (1988). *Educating reason: Rationality, critical thinking, and educa-*

tion. New York: Routledge Press.
Snow, M. & Brinton, D. (1988). Content-based language instruction: Investigating the effectiveness of the adjunct model. *TESOL Quarterly, 22*, 533-574.
Tarvin, W., & Al-arishi, A. (1991). Rethinking communicative language teaching: Reflection and the EFL classroom. *TESOL Quarterly, 25*, 9-27.
Walsh, D., & Paul, R. (n.d.). *The goal of critical thinking: From educational ideal to educational reality*. American Federation of Teachers Educational Issues Department.

(Received July 3, 1996; revised Nov. 20, 1996)

Appendix: The Moorburg Letter (Introduction and paragraphs 3, 6, and 8)

Dear Editor:

Overnight parking on all streets in Moorburg should be eliminated. To achieve this goal, parking should be prohibited from 2 a.m. to 6 a.m. There are a number of reasons why any intelligent citizen should agree.

3. Traffic on some streets is also bad in the morning when factory workers are on their way to the 6 a.m. shift. If there were no cars parked on these streets between 2 a.m. and 6 a.m., then there would be more room for this traffic.

6. Last month, the Chief of Police, Burgess Jones, ran an experiment which proves that parking should be prohibited from 2 a.m. to 6 a.m. On one of our busiest streets, Marquand Avenue, he placed experimental signs for one day. The signs prohibited parking from 2 a.m. to 6 a.m. During the four-hour period, there was *not one accident* [italics added] on Marquand. Everyone knows, of course, that there have been over four hundred accidents on Marquand during the past year.

8. Finally, let me point out that the Director of the National Traffic Safety Council, Kenneth O. Taylor, has strongly recommended that overnight street parking be prevented on busy streets in cities the size of Moorburg. The National Association of Police Chiefs has made the same recommendation. Both suggest that prohibiting parking from 2 a.m. to 6 a.m. is the best way to prevent overnight parking.

Sincerely,
Robert R. Raywift

(Ennis & Weir, 1985, p. 13)

Contrastive Rhetoric in Letter Writing: The Interaction of Linguistic Proficiency and Cultural Awareness

Taeko Kamimura
Senshu University

Kyoko Oi
Toyo Gakuen University

This study examines the relationship between Japanese college students' proficiency in English and their cultural awareness toward a target-language culture (North American). Japanese EFL students were divided into four groups according to high and low English proficiency and high and low cultural awareness based on performance on the CELT and the researchers' cultural awareness (CA) test. The four groups wrote college and scholarship application letters. The letters were analyzed in terms of frequency and content of different Semantic Formulas. The results show that subjects with both high English proficiency and high cultural awareness manifested the rhetorical patterns closest to those in the native speakers' English letters of application. Subjects with low English proficiency and low cultural awareness showed the rhetorical patterns closest to those in the Japanese letters of application. Subjects who lacked either the sufficient level of English proficiency or cultural awareness produced writing which varied from the target style. Results indicate that cultural awareness may be as important an element as English ability in student writing.

本研究は日本の大学生の英語能力と目標言語（アメリカ英語）の文化に対する気づきの間の関係を考察するものである。日本の大学生4グループが大学入学申込および奨学金申請の手紙を書いた。被験者は、CELTによって測定された英語能力と筆者の作成した文化的気づきを測定するテストの結果に従ってグループ分けされた。手紙は意味公式の頻度と内容という観点から分析された。その結果、英語能力が高く文化的気づきも高い被験者は母語話者の書く同様の手紙に最も近い修辞パターンをしめした。英語能力も文化的気づきも低い被験者は、日本語で同様の目的の手紙を書く場合に最も近い修辞パターンを示した。英語能力または文化的気づきのどちらかが欠ける被験者は、目標のスタイルとは異なる手紙を書いた。これらの結果は、文化的気づきが学生の英作文能力の重要な要因であるかもしれないことを示している。

JALT Journal, Vol. 19, No. 1, May, 1997

Robert Kaplan first proposed the notion of contrastive rhetoric in 1966. Since then many researchers of writing across the world have engaged in active research in this field. Enough evidence has been reported to support Kaplan's claim that each language has a culturally-preferred way of organizing ideas in discourse (e.g., a linear development in English vs. an indirect approach in a gyre style in Oriental languages), and that writers from different linguistic-cultural backgrounds transfer their preferred discourse patterns when they write in other languages. Among those researchers who have dealt with the differences of Japanese and English are Hinds (1979, 1980, 1981a, 1981b, 1987), Connor and McCagg (1983, 1987), Kobayashi (1984), Oi (1986), Mok (1993), Fisher-Stoga (1993, 1995), and Kimball (1996). All of them, however, have dealt with expository or argumentative writings in academic settings. Research oriented toward more pragmatic, non-academic perspectives has been scarce; only Jenkins and Hinds (1987) discusses the rhetorical differences in a more pragmatic context, namely business letter writing.The present study involves the comparison of rhetoric in letter writing in English and Japanese. This study is enforced by the following observations: 1) letter writing (such as social letters and business letters) is important to Japanese EFL students since it is the type of English writing that the students will face most frequently; 2) among several kinds of letter writing, however, the one they face immediately would be a letter of application because of an increasing number of Japanese students wish to study in American colleges, and 3) letter writing carries a pragmatic function to convey an intended meaning to a specific target audience in a particular culture.

The scheme of the present study is derived from the findings of two previous studies. The first study (Oi & Sato, 1990) investigated whether rhetorical transfer would be observed in Japanese EFL students' letter writing, comparing their letters with those of native speakers of English. The comparison was threefold. The first group was composed of Japanese students writing in Japanese, to determine the nature of Japanese rhetoric; the second group was composed of Japanese students writing in English; and the third group was native speakers of English writing in English. They all wrote under the same directions with the same assignment, though the directions were given in different languages depending on the groups: i.e., the first group in Japanese and the second and third groups in English. In this research, not only did we find rhetorical differences in letter writing between Japanese and American writers, we also found noticeable evidence of rhetorical transfer by Japanese students. However, this research did not take into consideration the En-

glish proficiency levels of Japanese students, so whether or not the difference was merely due to developmental factors was unknown.

The second study (Sato & Oi, 1990) was conducted under a similar format. However, this time the Japanese students were divided into two groups, high and low, according to level of English proficiency. The study showed rhetorical transfer was observed across the proficiency levels, indicating that the presence or absence of students' rhetorical transfer is not determined by English proficiency level alone. That is, there must be other factors.

The Study

From our previous studies we found that English proficiency alone is not a decisive factor in affecting Japanese EFL students' writing behavior in letter writing. Therefore, in the present research, we introduced a new factor, cultural awareness, which is in the domain of pragmatic competence. We define cultural awareness as one's familiarity with the perceptive and behavioral patterns in a target culture, American culture in this study. Combining these two factors, we propose the following research questions:

1) Does the degree of cultural awareness affect the Japanese EFL students' writing behavior in letter writing?

2) If the degree of cultural awareness is related to the students' writing behavior, what are the roles of cultural awareness and English proficiency respectively in the students' letter writing?

Method

Subjects: Subjects (N=42) were selected from students enrolled in three college EFL writing classes. The Ss were sophomore English majors at a Japanese college. They had practiced some narrative and expository writings, but they had experienced no formal training in letter writing. The Ss included in the study were selected, following the procedure outlined below.

Procedure

Task: All students in the three classes were asked to write a letter of application to a college in English. They were told to read a notice which announced the offering of a scholarship at a college in America and to write a letter applying for the college and the scholarship.

Instruments: Based on the results of the two former studies, we introduced two new factors for consideration. One is the variation in English proficiency levels of Japanese EFL students, and the other is the degree of their cultural awareness of American culture.

In order to measure the English proficiency levels, we administered the CELT (Comprehensive English Language Test) (Harris & Palmer, 1986). To test the degree of cultural awareness, we devised our own cultural awareness (CA) test. For the cultural awareness test, we selected 20 questions regarding critical situations that reflect crucial differences between Japanese culture and American culture (see Appendix). These questions covered the social, school, and workplace environments. The questions were devised by us, but the ideas were extracted from various books about cross-cultural communication that emphasize the differences between Japanese culture and American culture (e.g., Condon & Yousef, 1975; Sakamoto & Naotsuka, 1982; Barnlund, 1975, 1989; Furuta, 1987, 1990; Nishida, H., 1989; Nishida, T., Nishida, H., Tsuda, & Mizuta, 1989; Sherard, 1989; Matsumoto, 1994). The CELT and CA were administered to all students in the three classes. Based on the scores, we divided the students into the following four groups, W, X, Y and Z (Table 1). The W Group included students who scored high in both the CELT and CA test; the X Group scored high in the CELT but low in the CA test; the Y Group scored low in the CELT but high in the CA test; the Z Group scored low on the both tests. The demarcation line for "high" was placed at the upper 30 percent line among all the students and "low" at the lower 30 percent line for both tests. That was 165 points for high and 140 points for low on the CELT, and 15 points for high and 11 points for low on the 20-point CA test. Students who fell into these categories were selected as the subjects (N=42). Other students were excluded from the study.

Table 1: Classification of Subjects

CELT	CA	
	High ≥ 15	Low < 11
High ≥ 165	W (n = 9)	X (n = 9)
Low < 140	Y (n = 8)	Z (n = 16)

W: CELT-High, CA-High; X: CELT-High, CA-Low; Y: CELT-Low, CA-High; Z: CELT-Low, CA-Low

Analysis

Each of the letters written by the Ss was segmented according to idea units called Semantic Formulas (SFs), as done in our previous two studies. This analytical measure was based on a study by Beebe, Takahashi, and Uliss-Weltz (1990), which dealt with the analysis of refusals in speech. The SFs employed in the present study and a sample of each SF are:

1. Identification (ID): *I am a student at — college.*
2. Social talk (ST): *How are you?*
3. Referring to the ad. (RE): *I saw your ad about the scholarship.*
4. Writing a letter (WR): *So I am writing an application letter.*
5. Application message (AP): *I decided to apply.*
6. Reason (REA): *I'm interested in American culture and to study in America.*
7. Qualification (QUA): *I have a 3.8 grade point average (on a 4.0 scale) at —, and a score in the upper 20% bracket on the SAT test.*
8. Disqualification (DIS): *I'm afraid of going to the U.S. by myself.*
9. Petition (PE): *Could you please help me?*
10. Personal appeal (PA): *With these experiences, I feel I could make a positive contribution to ABC College and hope you will consider my application.*
11. Reference (REF): *I am enclosing a reference from Mr. Kempski, Head of the History department.*
12. Promise (PR): *I'll study hard.*
13. Apology (AP): *I'm sorry I have a favor.*
14. Request for information (REQ): *Please send me any forms that need completing.*
15. Closing remark (CR): *I would appreciate for your kindness.*
16. Expecting a reply (EX): *I'm looking forward to your letter.*

We first analyzed each letter into a sequence of the different SFs (see "Analysis of Sample Writing" below for sample analyses). We further analyzed those SFs in two aspects: 1) frequency (how often they appeared) and 2) content (concrete examples of SFs). The results were compared with those obtained in the two earlier studies in respect to the nature of mother-tongue writing of Japanese colleges students writing in Japanese and the target-language writing of native speakers writing in English.

Results and Discussion

Frequency

Each application letter was segmented according to the 16 different SFs, and how often each SF appeared was examined across the four different groups. Table 2 shows the frequencies of the 16 SFs identified in the four groups as well as those identified in the American subjects writing in English and the Japanese subjects writing in Japanese in the Oi & Sato study (1990).

In previous studies (Oi & Sato, 1990; Sato & Oi, 1990), the following differences were found between the Japanese and American subjects' application letters:

Table 2: Frequency of the Semantic Formulas Used by Different Groups

SF	Group					
	American (n=13)	W (n=9)	X (n=9)	Y (n=8)	Z (n=16)	Japanese (n=30)
ID	11 (84.6%)	8 (88.9%)	6 (66.7%)	5 (62.5%)	10 (62.5%)	16 (53.3%)
ST	0 (0%)	0 (0%)	2 (22.2%)	1 (12.5%)	9 (56.3%)	8 (26.7%)
AD	6 (46.2%)	5 (55.6%)	4 (44.4%)	4 (50%)	8 (50%)	25 (83.3%)
WR	6 (46.2%)	0 (0%)	1 (11.1%)	1 (12.5%)	2 (12.5%)	6 (20%)
AM	10 (76.9%)	9 (100%)	6 (66.7%)	6 (75%)	11 (68.8%)	18 (60%)
REA	6 (46.2%)	9 (100%)	7 (77.8%)	7 (87.5%)	13 (81.3%)	30 (100%)
QUA	8 (61.5%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (3.3%)
DIS	0 (0%)	0 (0%)	2 (22.2%)	1 (12.5%)	3 (18.8%)	5 (16.7%)
PE	1 (7.7%)	0 (0%)	3 (33.3%)	2 (25%)	8 (50%)	13 (43.3%)
PA	5 (38.5%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
REF	3 (23.1%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
PR	0 (0%)	1 (11.1%)	2 (22.2%)	3 (37.5%)	6 (37.5%)	13 (43.3%)
AP	0 (0%)	0 (0%)	2 (22.2%)	0 (0%)	0 (0%)	0 (0%)
IM	7 (53.8%)	5 (55.6%)	2 (22.2%)	2 (25%)	5 (31.3%)	7 (23.3%)
CR	7 (53.8%)	3 (33.3%)	1 (11.1%)	0 (0%)	2 (12.5%)	14 (46.7%)
EX	2 (15.4%)	3 (33.3%)	2 (22.2%)	2 (25%)	3 (18.8%)	2 (6.7%)

ID=Identification; ST=Social talk; RE=Referring to the ad; WR=Writing a letter; AP=Apllication message; REA=Reason; QUA=Qualifier; DIS=Disqualification; PE=Petition; PA=Personal appeal; REF=Reference; PR=Promise; AP=Apology; REQ=Request; CR=Closing remark; EX= Expecting a reply

1) Many American subjects used *persuasive* strategies, showing their abilities and previous experiences with concrete data. They often referred to their academic records and underscored their qualifications.

2) The Japanese subjects, regardless of the language used (English or Japanese) or of their English proficiency levels, tended to use *emotional* strategies meant to attract the reader's sympathy.

The Japanese subjects realized these *emotional* strategies by using such SFs as "social talk," "disqualification," "petition," and "promise." Those four SFs were rarely found in the letters by the American subjects. Our previous studies suggested that these four formulas were evidence of the Japanese subjects' negative transfer from Japanese writing and that these formulas were culturally inappropriate in English letter writing when addressed to American readers.

The SF "social talk" is a clear transfer of the Japanese usual letter format in which one is supposed to open a letter message with either greetings such as *I trust this finds you in good health*, or references to the weather such as *It has been awfully hot this summer. How are you coping with the heat?* It is considered impolite just to start business abruptly in any Japanese letter, even in a business letter. The Japanese students are so much used to this writing convention that they tend to persistently transfer this SF negatively in English application letters.

The SF "petition" is a desperate-sounding plea such as *Please give me a chance.* This type of pathetic tone is unfitting in an English letter of application.

The SF "disqualification" is a statement like *My English is not good.* Judging from the American practice of a letter of application to a college, this statement is like a taboo, admitting a reason for disqualification as an applicant. We interpret this as a reflection of the Japanese tendency to value modesty or understatement of oneself as often described in many studies (among them, Condon & Yousef, 1975). Concerning the Japanese tendency to resort to "petition" and "disqualification" when writing a letter of application, an exemplary anecdote is introduced in Sakamoto and Naotsuka (1982, pp. 34-37).

A typical example of the SF "promise" is a statement such as *I'll study hard if I'm admitted to your school.* A number of Japanese students finish their letters with "promise" following "petition." We regard it as a reflection of the Japanese formulaic expression *gambarimasu*, which can be translated as *I'll do my best.*

The present study, therefore, focuses on those four SFs (social talk, disqualification, petition, and promise) particularly identified as typical in Japanese Ss' letters of application.

Table 3 shows the frequencies (in percentage) of the four SFs used by the different groups (W, X, Y, Z) as well as those used by the American subjects writing in English and the Japanese subjects writing in Japanese (Oi & Sato, 1990).

Table 3: Frequencies of the Four Sematic Formulas (%)

SF (Sematic Formulas)			Group			
	American	W	X	Y	Z	Japanese
ST (Social talk)	0	0	22.2	12.5	56.3	26.7
DIS (Disqualification)	0	0	22.2	12.5	18.8	16.7
PE (Petition)	7.7	0	33.3	25	50	43.3
PR (Promise)	0	11.1	22.2	37.5	37.5	43.3

Each of the four groups demonstrated unique writing patterns. The writing patterns of the W Group, high English proficiency and high cultural awareness, were most similar to the patterns found for American subjects in the previous studies. No subjects in the W Group employed "social talk," "disqualification," or "petition," and only one subject included a "promise" (11.1%).

In contrast, the Z Group, low English and low cultural awareness, varied most from the native English speakers' pattern and showed the clearest transfer from Japanese. The Z Group ranked highest in the frequencies of three SFs: "social talk" (56.3%), "disqualification" (18.8%), and "petition" (50%), and second highest in the frequency of "promise" (37.5%).

The X Group, high English proficiency and low cultural awareness, and the Y Group, low English proficiency and high cultural awareness, were between the W and Z Groups. Though the differences between the X and Y Groups were not large, the percentages of the three out of four Japanese-oriented SFs were higher in the X Group (22.2% for "social talk," 22.2% for "disqualification," and 33.3 % for "petition") than in the Y Group (12.5% for "social talk," 12.5% for "disqualification," and 25% for "petition"). This suggests that both the writing patterns of X and Y Groups differed less from the target pattern than the Z Group, but the two groups did not as closely approximate the target style as the W Group. Additionally, the X Group demonstrated a greater variation from the target pattern than the Y Group.

Content

The next step of analysis was to examine what the subjects actually wrote for each of the four SFs in the letters of application. For each SF, several subcategories were established to examine the actual content of the four SFs in detail. For example, the SF "social talk" was further classified into such subcategories as *Hello*, *How do you?* and *How are you?* Table 4 illustrates the results of the content analysis.

The content written by the Z Group varied most from the pattern for the English letter of application and reflected most clearly the practice of the Japanese letter of application. The typical pattern of a letter of application by the Z Group started with a colloquial social talk *Hello* (31.3%), said, *I don't have enough English ability* (12.5%), yet pleaded for the scholarship by saying, *Please give me the scholarship* (25%), and ended with a promise by saying, *I will study hard* (31.3%). Thus, the letters by the Z Group had emotional and pathetic tones. In contrast, there was only one example of "promise" for the W Group: *I will make efforts* (11.1%). Emotional and pathetic tones were not perceived in the letters by the W Group.

The X and Y Groups manifested only one or two examples of some subcategories of the four SFs. The general tendency, however, was for both the X and Y Groups to be positioned between the W and Z Groups, as was found in the results in the frequency count of the SFs. The letters by the X and Y Groups did not sound as pathetic and emotional as those by the Z Group, but they were not as completely free of these tones as those by the W Group. Furthermore, such emotional and pathetic tones were somewhat stronger in the letters of the X Group than those of the Y Group.

Analysis of Sample Writing

Representative samples of the letters done by Ss from each of the four groups, W, X, Y, and Z, help explain the characteristic writing patterns. Each sample letter is analyzed by SFs, with errors left intact.

Sample 1, written by S-1 in the W Group (high English proficiency and high cultural awareness), is quite close to the target letters of application by the native speakers of English. The letter concisely conveys the intended message by including such SFs as "identification," "reason," "application message," "request for information," and "closing remark." None of the four SFs which characterize Japanese letters of application (social talk, disqualification, petition, and promise) are included.

Table 4: Content of the Four Semantic Formulas

SF Subcategory	Group					
	American(n=13)	W (n=9)	X (n=9)	Y (n=8)	Z (n=16)	Japanese(n=30)
Social Talk						
Hello	0 (0%)	0 (0%)	1 (11.1%)	0 (0%)	5 (31.3%)	2 (6.7%)
How do you do?	0 (0%)	0 (0%)	1 (11.1%)	1 (12.5%)	2 (12.5%)	6 (20%)
How are you?	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (12.5%)	0 (0%)
Disqualification						
Don't know reality	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (3.3%)
Don't have enough English ability	0 (0%)	0 (0%)	1 (22.2%)	1 (12.5%)	2 (12.5%)	4 (13.3%)
Afraid of going to America	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (3.3%)
Cannot express opinions	0 (0%)	0 (0%)	1 (11.1%)	0 (0%)	1 (6.3%)	0 (0%)
Petition						
Help me	0 (0%)	0 (0%)	1 (11.1%)	0 (0%)	0 (0%)	2 (6.7%)
Admit me	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (18.8%)	2 (6.7%)
Understand me	0 (0%)	0 (0%)	1 (11.1%)	0 (0%)	0 (0%)	7 (23.3%)
Give me the scholarship	0 (0%)	0 (0%)	1 (11.1%)	1 (12.5%)	4 (25%)	1 (3.3%)
Give me a good answer	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (3.3%)
Give me a chance	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (6.3%)	0 (0%)
Promise						
Study hard	0 (0%)	0 (0%)	1 (11.1%)	2 (25%)	5 (31.3%)	5 (16.7%)
Make efforts	0 (0%)	1 (11.1%)	0 (0%)	0 (0%)	0 (0%)	3 (10%)
Lead a full life	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (6.7%)
Make good use of the scholarship	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (10%)
Get something in America	0 (0%)	0 (0%)	1 (11.1%)	0 (0%)	0 (0%)	0 (0%)
Do my best	0 (0%)	0 (0%)	0 (0%)	1 (12.5%)	0 (0%)	0 (0%)
Mature	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (6.3%)	0 (0%)

Sample 1:
S-1, W Group, High English Proficiency and High Cultural Awareness

Dear Mr. Thompson:

*IDENTIFICATION
[I am a student in _____ University in Japan and I'm very interested in your exchange program between ABC College and our university.]

*REASON
[I'm studying English and American Literature here and I believe studying in your college will much help my research in as well as improving my English skills.]

*APPLICATION MESSAGE
[Therefore I do want to apply for this program. And I would like to apply for the scholarship you offer simultaneously.]

*REQUEST FOR INFORMATION
[Please send me a brochure and/or more details about them.]

*CLOSING REMARK
[Your most up-to-date information will be tremendously appreciated.]

Sincerely yours,

Sample 2 was written by S-2 in the X Group (high English proficiency and low cultural awareness). Since the level of S-2's English proficiency is high, there are no major grammatical mistakes. It does not include such Japanese-oriented SFs as "petition" and "promise." However, S-2 employs "disqualification" and underestimates self-worth, in contrast with the practice in the English letter of application where the use of "persuasive" strategies is expected. For instance, S-2 writes, *I can neither understand what English speakers say nor express my thought in English well*, and *I'm not good at express my opinion to other people*. Thus, even though S-2 exhibits no substantial grammatical mistake in the domain in linguistic competence, problems in cultural awareness in the realm of pragmatic competence are exhibited.

Sample 2:
S-2, X Group, High English Proficiency and Low Cultural Awareness

*REASON

[I have studied English since I was a junior high school student. English is not so easy to learn, but I'm interested in it very much, because the pronunciation of English words is very different from one of Japanese words, and I like pronouncing them very much.]

*DISQUALIFICATION

[To my regret, Japanese education in English is not so good for learning English conversation. I don't think I have much trouble reading English, but I can neither understand what English speakers say nor express my thought in English well.]

*REASON

[Living in the United States is the best way to improve such troubles of mine. That is the first reason I want to study at your college. The second one is that I want to take part in an active lessons.]

*DISQUALIFICATION

[I'm not good at express my opinion to other people.]

*REASON

[It' OK in Japan, but it cannot be allowed in other countries, so I want to train myself in active discussions in lessons at a college in the United States.

I have been to the United States to learn English before. My parents paid for me in that case. I thank them very much, but I can't have them pay any more.]

*APPLICATION MESSAGE

[So I'd like to be offered scholarship.]

Sample 3, from S-3, is representative of the Y Group (low English proficiency and high cultural awareness). S-3 does not include any of the four Japanese-oriented SFs. However, S-3's lack of organization makes the style far different from that found in the target letters of application. Reasons for application are mentioned sporadically and the self-introduction is begun, in an inappropriate place, too abruptly.

Like Sample 2, Sample 3 is not an appropriate English letter of application, but for different reasons. Although Sample 3 has no critical problem in content from a pragmatic aspect, it has many problems in grammar and organization.

Sample 3:
Y Group, S-3, Low English Proficiency and High Cultural Awareness

Dear Mr. Thompson,

* APPLICATION MESSAGE
[I want to get the scholarship.]

*REASON
[Because My father has been sick since last year. And My family is very poor. I can't afford money to go the college.]

*APPLICATION MESSAGE
[But I want to study English and literature in ABC college. I want to go to America.]

*REASON
[I'm interested in American costums, culture, family life and eating life. And I want to understand American people and watch beauties of nature, town.]

*IDENTIFICATION
[Introduction my self and my family.
My name is _____. I'm nineteen years old. I'm _____ University college student. My hobbies are playing tennis, watching movies, cooking, and shopping. I have four members. My father, My mother, my brother, me.
My father is 57 and businessman. But he is sick now. My mother is 48 and House wife. My brother is 23 and he graduated University this spring but he doesn't catch a job and in house he is studying law everyday for exercise.]

*REASON
[I want to learn Literature and American life. So I want to speak English.]

Sample 4, by S-4, represents the Z Group (low English proficiency and low cultural awareness). Sample 4 is linguistically unacceptable, with many errors in sentence construction and no organization as a paragraph. S-4 includes three of the four Japanese-oriented SFs: "social talk," "promise," and "petition." The letter begins with *hello*, makes several promises, saying, *If I go to ABC College, I study harder than now*, and *I will grow than now and I will come back to Japan!* In addition, S-4 petitions Mr. Thompson, saying, *Mr. Thompson, I want to know American people and culture. Please get the chance to me.* Thus, it is inappropriate as an English letter of application, both linguistically and culturally.

Sample 4:
S-4, Z Group, Low English Proficiency and Low Cultural Awareness

Dear Mr. Thompson:

*SOCIAL TALK *IDENTIFICATION
[Hello,] [My name is _____. I am twenty years old now. I am interested in American people and culture. But I've never seen foreign countries. I want to go to American very much. Of course, I am studying hard very day.]

*APPLICATION MESSAGE *REASON
[I want to get the scholarship.] [Because to help my home's life. My brother is high school student and my home is very new. Going to America need much money.]

*PROMISE
[If I go to ABC College, I study harder than now.]

*REASON
[And I want to make many foreign friends there. I think American is very friendly and kindly. Sure, I will get nice relationship with them.]

*PETITION
[Mr. Thompson, I want to know American people and culture. Please get the chance to me.]

*PROMISE
[After year I will grow than now and I will come back to Japan!]

Conclusion

In this study, we have examined the rhetorical differences between Japanese and English letter writing. The Japanese EFL students in this study seem to transfer Japanese rhetorical patterns into English when they write in English. Of immediate concern to the present study was examination of two factors, English proficiency and cultural awareness, which may determine the degree of rhetorical transfer.

The first research question sought to determine whether the degree of cultural awareness affects Japanese EFL students' letter writing behavior. The results of the present study answer this question affirmatively: both the level of English proficiency and the degree of cultural awareness of the target culture affect Japanese EFL students' letter writing behavior. The second research question concerned the respective roles of cultural awareness and English proficiency in the Japanese EFL students' writing performance. The results show the following tendencies:

1. Students with high English proficiency and high cultural awareness produce letters closest in style to that of native speakers of English;
2. Students with low English proficiency and low cultural awareness produce letters closest in style to that of native speakers of Japanese;
3. Students with high English proficiency and low cultural awareness produce letters with culturally inappropriate content but acceptable English;
4. Students with low English proficiency and high cultural awareness produce letters with generally culturally appropriate content but problematic English.

In order for our EFL students to compose in a way which is acceptable to an English-speaking audience, we need to develop not only their English proficiency but also their cultural awareness. This is especially so in letter writing, which carries a more pragmatic function than writing such as exposition and argumentation. As Mok (1993) asserts:

> Awareness of the [cross-cultural] differences is important because it makes students realize that to become part of the target language discourse community, they need to develop new attitudes, to meet certain criteria of the target language's traditions, and, in some cases, to put aside their native language habits. (p. 157)

We need, therefore, to develop teaching methods and teaching materials which integrate cultural factors with linguistic ones. Questions such as those used in our cultural awareness test could be modified and turned into instructional tasks. Alternatively, students could conduct their own SF analyses: first writing application letters, then analyzing the SFs in the draft to see whether L1-based SFs were included, and finally revising the letters into acceptable English letters of application.

More research will be needed to determine whether cultural awareness is a critical factor when the Japanese EFL students engage in other types of letter writing. In addition, research will be needed to examine whether the types of cultural and rhetorical instruction suggested above have positive effects on EFL students' writing performance.

Acknowledgments

We are grateful to the editor and two anonymous JALT Journal *reviewers for their careful reading and valuable comments.*

Taeko Kamimura teaches EFL writing and applied linguistics at Senshu University. Her interests include the relationship between L1 and L2 composing processes.

Kyoko Oi teachers EFL writing and applied linguistics at Toyo Gakuen University. Her research interests include the relationship between language and culture.

References

Barnlund, D. C. (1975). *Public and private self in Japan and the United States.* Tokyo: Simul Press.

Barnlund, D. C. (1989). *Communicative styles of Japanese and Americans.* Belmont, CA: Wadsworth Publishing Company.

Beebe, L, Takahashi, T., & Uliss-Weltz, R. (1990). Pragmatic transfer in ESL refusals. In R. C. Scarcella, E S. Anderson, & S. D. Krashen (Eds.), *Developing communicative competence in a second language* (pp. 55-73). New York: Newbury House.

Condon, J. C., & Yousef, F. (1975). *An introduction to intercultural communication.* New York: Macmillan Publishing Company.

Connor, U., & McCagg, P. (1983). Cross-cultural differences and perceived quality in written paragraphs of English expository prose. *Applied Linguistics, 41,* 259-268.

Connor, U., & McCagg, P. (1987). A contrastive study of English expository prose paragraphs. In U. Connor, & R. B. Kaplan (Eds.), *Writing across languages: Analysis of L2 text* (pp. 73-86). Reading, MA: Addison-Wesley.

Fisher-Stoga, F. (1993). Convention and composition in the Japanese ki-sho-ten-ketsu: Towards a methodology of contrastive rhetoric. *Tokyo Daigaku Gaikokugoka Kenkyukiyo (The Proceedings of the Department of Foreign Languages and Literature, College of Arts and Sciences, University of Tokyo), 41*(3), 130-168.

Fisher-Stoga, F. (1995). An independent learning: Lafcadio Hearn as educator. *Tokyo Daigaku Gaikokugoka Kenkyukiyo (The Proceedings of the Department of Foreign Languages and Literature, College of Arts and Sciences, University of Tokyo), 43*(3), 24-43.

Furuta, G. (Ed.) (1987). *Ibunka komyunikeshon (Intercultural communication).* Tokyo: Yuhikaku.

Furuta, G. (Ed.) (1990). *Ibunka komyunikeshon kii waado (Key words for intercultural communication).* Tokyo: Yuhikaku.

Harris, D. P., & Palmer, L. A. (1986). *CELT (A Comprehensive English Language Text for learners of English).* New York: McGraw-Hill Book Company.

Hinds, J. (1979). Organizational patterns in discourse. In T. Givon (Ed.), *Syntax and semantics 12: Discourse and syntax* (pp. 135-157). New York: Academic Press.

Hinds, J. (1980). Japanese conversation, discourse structure, and ellipsis. *Discourse Processes, 3,* 263-286.

Hinds, J. (1981a). *Ellipsis in Japanese.* Alberta: Linguistic Research.

Hinds, J. (1981b). Japanese expository prose. *Papers in Linguistics, 13*, 117-157.

Hinds, J. (1987). Reader versus writer responsibility: A new typology. In U. Connor & R. B. Kaplan (Eds.), *Writing across languages: Analysis of L2 text* (pp. 141-152). Reading, MA: Addison-Wesley.

Jenkins, S., & Hinds, J. (1987). Business letter writing: English, French, and Japanese. *TESOL Quarterly, 21*, 327-345.

Kaplan, R. (1966). Cultural thought patterns in inter-cultural education. *Language Learning, 16*, 1-20.

Kimball, J. (1996). Writing on academic topics: Externalizing rhetorical processes in an intercultural context. *JALT Journal, 18*, 55-66.

Kobayashi, H. (1984). Rhetorical patterns in English and Japanese. *TESOL Quarterly, 18*, 737-738.

Matsumoto, S. (1994). *Nihonbunka no tokushitsu (Characteristics of Japanese culture).* Tokyo: Kenkyusha.

Mok, W. E. (1993). Contrastive rhetoric and the Japanese writer of EFL. *JALT Journal, 15*, 151-161.

Nishida, H. (1989). *Jitsurei de miru nichibei communication gap (Communication gap between Japan and America examined by examples).* Tokyo: Taishukan.

Nishida, T., Nishida, H., Tsuda, Y., Mizuta, S. (1989). *Kokusai ningen kankeiron (Studies on international human relationships).* Tokyo: Seibunsha.

Oi, K. (1986). Cross-cultural differences in rhetorical pattering: A study of Japanese and English. *JACET Bulletin, 17*, 23-48.

Oi, K., & Sato, T. (1990). Cross-cultural rhetorical differences in letter writing: Refusal letter and application letter. *JACET Bulletin, 21*, 117-136.

Sakamoto, N., & Naotsuka, R. (1982). *Polite fictions.* Tokyo: Kinseido.

Sato, T., & Oi, K. (1990). *Nichiei retoriku no hikaku bunseki* (A contrastive analysis of Japanese and English Rhetoric), In JAAL-in-JACET (Ed.), *Oyoogengogaku no kenkyu (Studies in applied linguistics)* (pp. 146-155). Tokyo: Liber Press.

Sherard, M. (1989). *Tacit assumptions.* Tokyo: Macmillan Language House.

(Received March 1, 1996; revised September 2, 1996)

Appendix: Sample Questions from Cultural Awareness Test

Under the following circumstances, which behavior do you think you are likely to follow? Choose one from the two alternatives.

[Social life]

1. Six months have passed since you came to the U.S. with your parents. Today you have invited Mr. and Mrs. Brown to your home. They have been very kind to you since you came to the U.S. and this is their first visit. While chatting over tea, Mrs. Brown says to you, "Could I see the rest of your house?" To your regret, the house is far from being clean enough to show to other people. How would you react to Mrs. Brown's request?
   1) You would refuse, saying, "It's such a mess. I really cannot show you this time."
   2) You would show her around, saying, "It's a mess, but if you don't mind that."

2. You are asked to have tea by an American woman, Mrs. Anderson. It is hot, and you are thirsty. When Mrs. Anderson says, "What would you like? Would you like something hot or cold?" what would you say to her?
   1) You don't want to bother Mrs. Anderson, so you would say, "Anything will be fine with me."
   2) You would say clearly what you would like to have.

[School life]

3. You are now studying at D University in the U.S. and taking Sociology I among other courses. Professor Samson, who is teaching Sociology I, takes a discussion style in his class. Since your English is still not good enough, you cannot quite participate in the discussion with American students, although you are trying to. There is another Japanese student, Mariko, in this class. She is always quiet and does not contribute to the discussion. You came here two years earlier than Mariko. How would you advise her?
   1) You would suggest that she tell Professor Samson of her linguistic disadvantage and ask him to acknowledge her willingness to participate.
   2) You would advise her to participate in the discussion as actively and assertively as she can, seeking the professor's help after class as needed.

4. It has been a month since you began studying at B University in the U.S. The other day you were asked to give a speech for an audience comprised of professors. Although you are not confident of your English as it has been just a month, you've decided to give the speech. How would you deliver the speech?
   1) You would try to be confident of your English and not mention anything concerning the ability of your English.
   2) You would, first of all, tell the audience that your English is not good because you are afraid that the audience will be surprised at your poor English.

[Workplace]

5. You are employed by an American company. Yesterday you saw Jane, who is a co-worker, step into the elevator before Mr. Black, who is her boss. You are older than Jane. How would you feel about her behavior?

1) You would assume it natural since she is a woman.
2) You would try to reprimand her as you think she was being rude.

6. After graduation from college, you climbed up the ladder of success and are now a branch-office manager. As business is good this year, you are quite busy. Today you have work that needs to be done by tomorrow. Unfortunately tomorrow is Sunday. If you fail with this, it means a loss to the company, so you want your employees to come to work tomorrow. What would you do?

1) You would ask your employees to come to work even on Sunday, explaining to them it is for the sake of the company.
2) You would ask for volunteers to help with the project, stressing extra benefits for those who choose to do so.

実際使用アクティビティーを使用した海外での日本語教育例

加藤好崇

東海大学留学生センター

ネウストプニー（１９９５）によるとインターアクション教育のためには、「解釈アクティビティー」、「練習アクティビティー」、「実際使用アクティビティー」の三つのアクティビティーの適用が必要である。まず、「解釈アクティビティー」で講義などによる説明で知識を与え、「練習アクティビティー」でロールプレーなどの形で練習をする。そして、最後に「実際使用アクティビティー」において実際使用場面を教育過程の中に導入する。本稿は、上記の考え方をハバロフスク国立教育大学日本語学科において適用した授業報告例である。ここでは、日本人ビジネスマンへのインタビュー依頼の手紙、電話での場所や日時等の決定、インタビュー、結果発表などが「実際使用アクティビティー」として導入されている。また、それぞれの段階において「解釈」と「練習」のアクティビティーも「実際使用アクティビティー」に先立って実施された。最後にこのプロジェクトワークの結果得られた言語学上、社会言語学上、社会文化上の有効点、問題点を考察する。

Neustupny (1995) claims that learners need to be engaged in three kinds of activities to acquire interactive competence in a second language. The first of these is interpretation activities. These aim to provide learners with linguistic knowledge, including explanations by lectures and the learners' own research in the form of project work. Second is practice activities, which entail drills, role plays, and simulations. Third is performance activities. These differ from practice activities in that they provide authentic situations for learns to interact in the target language. Examples of performance activities are immersion courses and inviting native speakers of the TL to the classroom. This article reports the process and outcome of project work organized by two classes of advanced learners of Japanese at a Russian university following the framework suggested by Neustupny. This project focused on opinions of Japanese native speakers (NSs) working in Khabarovsk about Japanese and Russian working customs. The project, conducted during 10 classes of 80-minute each, followed five stages: preliminary preparation stage, writing a letter asking for an interview, making an appointment on the phone, interviewing a Japanese NS, and presentation of the results to the entire class. Each stage consisted of a sequence of interpretation activities, practice activities, and performance activities.

Interviews were recorded and used for follow-up interviews with the subjects. In addition, the author interviewed four of the eight Japanese NSs who had been interviewed by the subjects. From the analyses of the recorded NS/subject interviews, the follow-up interviews with subjects and NSs, and the content of the presentations, it was found that: 1) some of idioms and vocabulary encountered by the subjects for the first time during their interviews of the NSs were adopted into the subjects' active vocabulary; 2) some of the vocabulary provided by the instructor in the preliminary stage was reinforced through use in the interviews and presentations, again entering the subjects' active vocabulary; 3) subjects became aware of the variety of Japanese dialects; 4) subjects became aware of the Japanese view toward Russian customs, and 5) subjects participated more actively and spontaneously during the project than they did during classroom-based activities. Major problems perceived were: 1) the use of Japanese dialects by the NSs made comprehension difficult for the subjects; 2) the interviews were not always well structured and some of the Japanese NSs interviewed felt that they had been asked an unorganized list of questions, and 3) the content and the language of the answers by the Japanese NSs were not adjusted to the linguistic level of the subjects, who were then unable to use appropriate communication strategies to repair communication breakdowns. The author suggests similar projects be undertaken, especially ones with follow-up interviews of all the NSs interviewed, in order to establish clear criteria for teaching and evaluation. The author also suggests teacher intervention during the planning stage, especially in the preparation of interview questions, to aid learners in organizing the interview and keeping the language from the NSs at a level the learners would find manageable.

海外で日本語を学ぶ学生にとって、教師以外の日本人とのコミュニケーションは非常に貴重な体験となる。またその際、情報の伝達を目的としていく過程で、学生の目標言語を使用する能力も促進されると考えられる（Ellis, 1990)。本稿ではこうした考え方に基づき、筆者が勤務したロシア連邦ハバロフスク国立教育大学において試みられた授業例を紹介したい。この授業では、在住日本人へのインタビュー依頼の手紙、電話でのアポイントメント、実際のインタビューなどが主要なタスクとして組み込まれた。

この授業の具体的な目的は以下の3点である。

① 学生に日本人との実際のインターアクションを体験させ、またその中で日本語使用能力の向上を計る。

② 学生の学習意欲を高める。

③ 学生に日本語の中に存在する方言などの多様性や日本人の考え方の多様性を認識させる。

タスクベースシラバス

タスクの捉え方は研究者によって多少の差があるが（Richards, Platt & Weber, 1986; Nunan, 1989)、基本的には、目標言語を理解あるいは産出する過程で何らかの課題を達成していくことを言う。そして、このタスクによって授業はよりコミュニカティブになる可能性を持つ。

Nunan (1989, pp. 59-63)はこのタスクを構成する要素の一つである活動を次のような観点から分類している。第一に、活動自体のauthenticityの程度、第二に、その活動がスキルを身につけることを目的としているものか、実際のコミュニケーションの中でスキルを使うことを目的としているものか。第三に、それが流暢さを目的としているものか、正確さを目的としているものか、の三点である。しかし、これらの活動が行なわれる環境に関しては、Nunanは主に教室内に中心を置いている。

それに対して、Strevens (1987)は教室外で行われる活動の重要性を指摘している。活動例としては、ペンパルを持つこと、目標言語話者のビジネスコミュニティーと関係を持つこと、メディアを利用するものなどが挙げられている。またStrevensは、教室外で行われる活動は学生が実際のインターアクションに教師を介さず接することができ、また活動を管理する責任が学生に移動することにより学習意欲が高まるなどの利点を挙げている。

ネウストプニー(1995)もまた教室外での活動の必要性を認め、それを「実際使用アクティビティー」と呼ばれる活動の範疇に含めている。ネウストプニー(1995, p.23)によると、「実際使用アクティビティー」はその前段階として、知識を与えることを目的とした「解釈アクティビティー」、次にその知識を使って練習を繰り返す「練習アクティビティー」の存在が必要であり、実際のインターアクションを目的とする「実際使用アクティビティー」は最後に行なわれなければならないとしている。ネウストプニー(1995, pp. 20-30)はこの三つの「アクティビティー」について以下のように説明している。

① 「解釈アクティビティー」: 最も伝統的、かつ広く行われている方法で、具体的には講義による説明や学習者がプロジェクトの形などで「調べ」たりするものなどが含まれる。

② 「練習アクティビティー」: インターアクションの練習のためのものである。このアクティビティーには、文法的ドリル、ロールプレー、シミュレーションなどが含まれる。しかし、これらはあくまで練習として行われるもので自然さの程度に問題が残り、さまざまな状況が予想される実際の場ですぐに効果があるかどうか疑問である。

③ 「実際使用アクティビティー」: 「練習アクティビティー」と違い、実際の目標言語の使用場面を教育現場に取り入れたものである。例えば、教室内にお客を呼んで学生に質問をさせたり、あるいは学校内のいろいろな場所を案内させる、日本料理店に行くなどの活動が考えられる。また、オーストラリアのモナシュ大学で行われているイマーションコース（尾崎・ネウストプニー,1986）などもその例として挙げらる。

本稿の授業例も、この三つの「アクティビティー」の流れに沿って実施された。

実施コースの概要

本稿で述べるインタビューを主とした一連の授業は1994年11月から12月までの10数回（行事などの都合でクラス間に多少の差が生じた）の授業（各8 0 分）で実施された。実施された学年は3年（12名）および4年（12名）である。このプロジェクトでは3名ずつが1つのグループを形成し行動をとった。

テーマは各学年とも「ハバロフスクで働く日本人（日本人の労働観）」であり、これは筆者によって選択された。実施コース全体の構成は図1のようになる。

図1：実施コースの構成

①「予備段階」　テーマに沿った予備的な知識を与える

↓

②「手紙を書く段階」*〈1解釈→2練習→3実際使用〉インタビューの依頼の手紙を書く

↓

③「電話でアポイントメントをとる段階」〈1解釈→2練習→3実際使用〉インタビューの日時を決定する

↓

④「インタビュー段階」〈1解釈→2練習→3実際使用〉約束場所に行き、インタビューを実施

↓

⑤「発表段階」〈1解釈→2練習→3実際使用〉　インタビューの結果をグループ毎に発表

*「解釈」は「解釈アクティビティー」、「練習」は「練習アクティビティー」、「実際使用」は「実際使用アクティビティー」を表わす。

また、実際のインタビューの会話は学生によって録音され、筆者がそれを文字化した後、発表内容との比較および分析を行った。さらに、筆者はプロジェクト終了後、各グループのメンバーをそれぞれ呼び、インタビューの録音を聞きながら、フォローアップインタビュー（ネウストプニー, 1981, pp. 36-38）を試みた。また、インタビューを受けた日本人のうち、4名にもフォローアップインタビューを行った。

最後に、評価に関しては、今回は「発表段階」に限り、デス・マスの一貫、発表内容の明瞭さ・掘り下げ方・面白さ、敬語、明瞭化への配慮、発音などの項目で評価を行った。

各段階の詳細

「予備段階」（3, 4回分の授業）

この段階で学生はテーマに関連する背景的知識を学習した。内容としては、現在の日本の労働状況、女性の労働も含めた問題点、就職状況、サラリーマンの日常などが取り上げられ、以下の教材を中心としてすすめられた。またこの際、聞き取り問題、読解問題などのタスクが学生に課せられた。

＜教材：①ビデオテープ「サラリーマンの休日」（国際交流基金）②「日本を話そ

う」（ジャパンタイムズ）③「総合日本語中級」（凡人社）④「朝日新聞で日本を読む」（くろしお出版）⑤ 朝日新聞などから関連記事を利用した生教材＞

次に上で得た知識をロシア人や個人の立場で考え直させるために、ディベート（3対3で、「人生には余暇が最も大切」グループ対「人生には仕事が最も大切」グループ）とレポート（ロシア人の労働観や労働状況、あるいは日本との比較についてで、これは宿題として提出させた）を課した。

また、この段階までに筆者はインタビューを受けてくれる在住日本人を探し出し、インタビューの許可を得ている。そして、各グループにインタビューの相手を自由選択させた。

「手紙を書く段階」

①「解釈アクティビティー」（1回半～2回分の授業）

ここでは、「手紙の書き方事典」（三省堂）を基礎とし、手紙の基本的な構成、特殊な言い回しの例、あるいは手紙のサンプルを見せるなどして説明を行った。今回は特に未見の人に出すインタビュー依頼の手紙なので、それに関連する項目を中心に説明した。

②「練習アクティビティー」（一回分程度の授業、宿題）

グループごとにインタビュー依頼の手紙を作成。出来上がったものを黒板に書かせ、筆者が教室内でチェックした。この際、手紙の構成や特別な表現、敬語に特に注意させた。宿題として、筆者が日本に帰って一年経過したという想定で日本の筆者に出す手紙の作成をさせた。

③「実際使用アクティビティー」

学生は教室内で訂正された「依頼の手紙」を宿題として書き直しをした後、提出。筆者の許可がおり次第、清書し郵送した。具体的な手紙の内容としては、自己紹介、事情説明とインタビューの趣旨及び依頼、後日電話にてインタビューを受けてもらえるかどうか確認をしたい、といったものが含まれている。次のページの例はその一例である。

「電話でアポイントメントをとる段階」（1,2回分の授業）

①「解釈アクティビティー」

まず、各グループごとに、電話で話すべき内容（自己紹介、インタビューの目的、インタビューを受けてもらえるかどうかの確認、日時、場所などを）を確認させ、その流れを簡単にロシア語でメモさせる。次に一人代表者を出し、筆者と電話でアポイントメントをとるロールプレーをし、それを録音する。この内容を全体で聞き、敬語の誤り、言い残したことはないかどうかのチェックなどを全員で確認していく。最後に筆者が作成した例文を与えて確認。この際、使用される敬語に特に注意させた。

②「練習アクティビティー」

筆者から渡された例文をペアーで役割を交代しながら読んでみる。次に、例文を基

例１　学生による手紙の一例（縦書きを横書きに直したもの）

拝啓　突然お手紙を差し上げます失礼をお許しください。私どもは、ハバロフスク国立教育大学の四年生の者でございます。

本日は、折り入ってお願いがございます。実は、今、当校では「ハバロフスクで働く日本人」というテーマで勉強をしております。つきましては、ぜひ岡田様にそういったテーマで御意見をうかがいたいと思っております。食事や仕事、また日本人とロシア人の労働観の違いなどの問題についていろいろ教えていただけないでしょうか。

十二月中旬にインタビューを予定しております。　改めて本状がお手元に届きます頃を見はからってご連絡を申し上げ、御意向をうかがう所存です。

ご多忙の折りから恐縮に存じますが、どうかよろしくおはからいくださいますようにお願い申し上げます。

右、取り急ぎインタビューのお願いまで。　敬具

一九九四年十二月四日

ポストニコフ
ブルロワ
チナナリョーワ

岡田一也様

本にしながら一人がインタビューを受ける側、一人がする側になり、配布されたお互いの異なる予定表を見ながら日時を調節し、会う場所を決めるロールプレーを行う。その後、いくつかのグループに代表で会話をさせた（例２参照）。

③「実際使用アクティビティー」

「練習アクティビティー」の２、３日後の授業時間外に学校の職員室から実際にインタビュー依頼の電話をかけさせる。その場には電話をかける代表者の他に筆者、他のメンバーなども立ち会う。電話の後、筆者は簡単に学生にコメント程度の評価をした。

「インタビュー段階」

①「解釈アクティビティー」と②「練習アクティビティー」

ここでは二つのアクティビティーが交互に行われた。

まず、インタビュー場面において予想される全体的な流れを示したプリントを渡し、次にそれぞれの段階で必要と思われる表現等の確認をしていった。具体的には、挨拶、自己紹介の仕方、考え・意見・感想の聞き方、相槌、言い淀みの種類や機能、話題の変え方、コミュニケーション問題が起きた時の処理の仕方（聞き返し、言い換えを求める、相手の言葉を自分の言葉で言い換える、相手の言葉を確認する）、話を終えるときの方法などを以下の教材を利用しながら説明、練習を繰り返していった。

例2　電話での参考会話例

（本人の場合）
A：もしもし、（会社or本人の名前）ですが。
B：あ、もしもし、私、ハバロフスク教育大学の（自分の名前）と申しますが、～さんをお願いできますでしょうか。
A：はい、私ですが。
B：あ、始めまして。
A：始めまして。

（本人でない場合）
C：もしもし、（会社or本人の名前）ですが。
B：あ、もしもし、私、ハバロフスク教育大学の（自分の名前）と申しますが、～さんをお願いできますでしょうか。
C：はい、少々お待ちください。
...
A：もしもし、お電話代わりました。A　です。
B：あ、私、ハバロフスク教育大学の（自分の名前）と申します。始めまして。
A：始めまして。

B：あのう、先日私どもが差し上げましたお手紙はお手元に届きましたでしょうか。
A：ああ、はい、いただいています。何かインタビューをなさりたいとか...。
B：はい、実は今度、日本語のクラスで、「ハバロフスクで働く日本人」というテーマで日本の方にインタビューをして、その結果を発表することになったんです。それで、～さんにお仕事の内容や、日本人とロシア人の労働感の違いなどについて三十分程質問をさせていただきたいと思いまして...。
A：ああ、いいですよ。喜んでいたしましょう。
B：あ、どうもありがとうございます。あのう、それで、インタビューの日時と場所なんですが、私どもは（　）の（　）時ごろ、そちらにうかがおうかと考えているのですが、Aさんのご都合はいかがでしょうか。
A：そうですねえ、その時間はちょっと...。（　）の（　）時ごろはどうですか。
B：ええと、ちょっとお待ちください。...あ、はい、結構です。それでは（　）の（　）時に、同じ教育大学のDとEというものと三人で参りますので、よろしくお願いいたします。
A：はい、わかりました。それじゃあ、お待ちしています。
B：はい、それではどうもありがとうございました。失礼いたします。
A：さようなら。

＜教材：①「インタビューで学ぶ日本語」（凡人社）②「ロールプレーで学ぶ会話1＆2」（凡人社）③「待遇表現」（ジャパンタイムズ）＞（例3参照）

「練習アクティビティー」では主にペアーで会話文を読ませたり、またドリル、インタビューのロールプレーなどを行った。

また、社会文化的な側面、例えばお辞儀の仕方や勧められるまで椅子に座らない、足を組まないなどの確認もした。

③「実際使用アクティビティー」

各グループが決められた日時、場所に出かけて行ってインタビューをした。また、インタビューの内容は録音された。

「発表段階」

①「解釈アクティビティー」と②「練習アクティビティー」

対象となった学生達には新学期開始時から、発表や司会を行う活動を定期的に導入

していたので、今回は特に改めて何も行わなかった。しかし、確認のために、司会の始めの言葉、インタビューを受けた人の紹介、発表者の紹介、質問に対する答え方、感想の述べ方などをプリントして渡し、目を通させた。

③「実際使用アクティビティー」

グループごとにインタビューの結果を報告した。発表時間は各グループ２５分程度で、３人のうち１人が司会者と発表者を兼ね、残りの学生は発表だけをした。また、他のグループの発表に対して必ず一つは質問をするように求めた。この発表内容は録音された。

効果

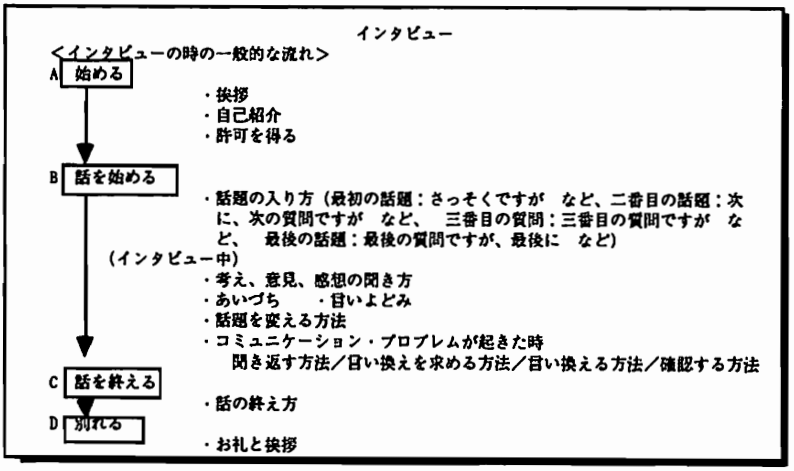
ここで挙げられる効果とは基本的にインタビューを行ったことによる学習効果のことで、インタビューとその発表内容、それにコース終了後の各学生に対するフォローアップインタビューの分析結果からわかったことである。

言語学的効果

この節では、インタビュー時、あるいはインタビューをきっかけにして、習得されたと思われる語彙、表現の例を挙げる。

習得されたかどうかの判断は、インタビュー後の筆者との自然な会話の中で適切にそれらを使用していた場合、「発表段階」でそれらの意味の説明を行ったり、正しい文脈の中で使用していた場合、また学生とのフォローアップインタビュー中に筆者にそれらの意味を適切に説明できた場合のみを考慮に入れた。

例３　インタビューの流れとそれぞれの段階の学習事項



また、習得されたと思われる語彙や表現が実際に学生にとって新しいものであったかどうかは、フォローアップインタビューの際の学生の証言によった。

①インタビュー中、会話の文脈から推測し、新しい表現の習得のきっかけを得た例。
[1]「郷に入っては郷に従え」
学生Aは、日本人がインタビュー中に使用したこの表現の意味を会話中に推測し、理解したと述べている。学生Aはインタビュー後、実際に筆者との会話の中でこの表現を数回適切に使用している。
[2]「日本人のお客はサービスにうるさい。」
この表現は学生Bにとって初めての表現だったが、文脈からその意味を推測し、その後発表までの準備の過程で完全に理解したと述べている。また、学生Bは発表の際にこの表現を適切に使用している。

②インタビュー中の日本人側から学生側への語彙知識の確認、あるいは学生側から日本人側への説明要求が日本人による語彙の説明を導き、その結果それらの語彙が習得されたと思われる例。学生はこれらの語彙の意味をフォローアップインタビューの際、筆者に適切に説明している。

語彙知識の確認の例：「保険、わかりますか？」、「営利活動ってわかる？」
説明要求の例　　：「責任感ってなんですか？」

③インタビュー中に「予備段階」で学習した語彙が現われ、習得が強化された例。以下の語彙も"と同様、筆者は学生とのフォローアップインタビューの中でそれぞれの語彙に対する学生の理解を確認している。

「残業」、「週休二日制」、「働き蜂」

④インタビュー中の意味不明の単語などを後に辞書で調べる過程で習得した例。
「楽天家」（この語を習得した学生B は、後に筆者との日常会話でも実際に使っている）、「利息」（この語を習得した学生C は発表時に他の学生からの質問に対して日本語で説明をした）

社会言語学的、社会文化的効果

①日本語に存在する多様性を知った。
青森県出身の会社員にインタビューを行った学生達は、その日本人のイントネーションが今までに接触した日本人と異なることに気付いたとフォローアップインタビューの際に語っている。また、この日本人は相槌の回数もかなり多く、同グループの一人の学生はこの相槌の個人的特徴にも言及している。

②日本人の視点から自分達の国を見ることができた。
以下の文は日本人によって発話されたものである。

「…ロシアの銀行はそういった意味で（日本の銀行と比べると）まだまだ、真似事…」
「ロシアでは、お店で袋がもらえませんね。」
「ロシアの旦那さんは奥さんの手伝いをしないんですか！」

上記のものは学生とのフォローアップインタビュー、あるいは後の筆者との会話の中で学生が特に印象を受けたと筆者によって判断されたものである。いずれも学生達にとって、日本人がロシアに関して考えていることや驚いたことを知ることができ、同時に、それらを通して逆に日本の様子や日本人の自国に対する考え方を学ぶことができた。

教育上の効果

Strevens (1987, p. 171)は、教室外での母国語話者との実際のインターアクションを授業に取り入れることによって、学生にとって授業が興味深いものとなる、学生が会話者としての役割を自覚する、学習意欲が向上するなどの効果がもたらされるとしている。また、Dickinson (1995, p. 174)によれば、学生に自らの学習に対する責任を与えることは、動機付けを促進し、学習を成功に導く条件の一つであるとしている。

今回のこのプロジェクトでも、学生達は教室外での実際のインターアクション場面を与えられ、また同時にこのインターアクション場面では教師の監督下を離れ、自ら言語使用の検証、修正を行っていく責任をも与えられた。このことは学生のより積極的、自発的な学習計画全体への参加態度を生んだようだ。また、マンネリ化した教室場面から離れることも学習意欲向上の一原因となったと思われる。

問題点

ここでの問題点とは5章同様、「インタビュー段階」を中心に筆者によって認識された問題点であり、学生へのフォローアップインタビュー、及びインタビューを受けた日本人側へのフォローアップインタビューを通して明らかになったものである。

言語学的問題点

インタビューの結果、間違った日本語を記憶してしまった学生がいた。

「るっか」、「えんたい、たいえん」

学生Eは発表時に意味不明の上記のような日本語を使用した。彼女はこれらを経済の専門用語であると理解していたが、フォローアップインタビューの際、一緒にテープを聞き直してみると、いずれも彼女が理解可能な語彙だった。つまり、「るっか」とは「物価」のことであり、「えんたい、たいえん」とは「円対ルーブル」の意であった。この場合、相手の日本人の方言（青森弁）や話し方が原因になったと思われる。

インタビューの技術上の問題点

①仕事に関する質問、生活に関する質問が交互に現われ、内容ごとに質問の順序が組み立てられておらず、インタビュー全体にまとまりが見られないグループがあった。この場合、日本人は矢継早にいろいろ質問を投げかけられているといった印象を持ったようだ。

②それぞれ分野の違う日本人に仕事に関する内容を聞く場合、当然その分野に関係する専門用語が出てくるはずである。したがって、それに対する各グループごとの事前の準備、または相手の答えの中に含まれる語彙あるいは内容の難度を下げるための質問作成段階でのコントロールが必要だった。あるグループは日本人の銀行家に「ハバロフスクの経済状態についてどう思われますか。」という質問をし、かなり難しい説明を受けた。そのため学生達にとってそのインタビューがほとんど内容のわからない、つまらないものとなってしまった。しかし、彼等は相槌だけは打ち続けたため、その銀行家はわかっているものと思い込み、ほとんど調整を加えずに話し続けた。この場合、説明要求などのストラテジーの使用も必要だったであろう。

③８グループ中３グループが録音に失敗した。録音は、後に評価をするための重要な材料となるものなので、録音機の状態を学生に前もってよく確認させておく必要があった。

その他

前述したようにインタビューを受けた日本人のうち４人に簡単なフォローアップインタビューを行ったが、すべての日本人についてさらに詳しく調査する必要があった。日本人側が不快に思ったり疲労感を覚えたりした状況を分析することは、次回の指導のポイントの有効な指標になりうる。また、これはインタビューだけではなく、手紙や電話の段階でも、できるだけ行なわれる必要があった。

おわりに

今回のプロジェクトは言語習得の促進を図り、学習に対する動機付けを強化し、さらに日本語や日本人の考え方の多様性を認識させるなどの点で一応の効果を達成したものと思われる。しかし、その分析対象が「インタビュー段階」の主に「実際使用アクティビティー」を中心としており、今後は各段階における３つのアクティビティーの流れが有機的にもたらす効果、また全段階の統合の結果による効果を詳しく検証する方法を考えていかなければならない。同時に、学生に対する評価に関しても「発表段階」以外の段階をどのように適切に評価していくかを検討していく必要がある。

また、自然なインターアクションの中で学生の日本語運用能力の向上を目指す場合、この様なプロジェクトを年に一度実施するだけではさほど効果は期待できない。プロジェクトは長期にわたるものである必要は必ずしもなく、短期的なものでも今回の様な「解釈アクティビティー」、「練習アクティビティー」、「実際使用アクティビティー」という流れに沿ったものを定期的にカリキュラムに組み込んでいく必要があろう。

加藤好崇はオーストラリア、モナシュ大学日本語学科マスターコース終了後、１９９４年６月より１９９５年６月まで国際交流基金、日本外交協会による派遣日本語講師としてロシア連邦ハバロフスク国立教育大学日本語科に在籍した。現在、東海大学留学生センターに勤務している。

参考文献

Dikinson, L. (1995). Autonomy and motivation: a literature review. *System, 23*(2), 165-174.

Ellis, R. (1990). *Instructed second language acquisition*. Oxford: Blackwell.

Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge: Cambridge University Press.

Richards, J., Platt, J., & H. Weber. (1986). *Longman dictionary of applied linguistics*. London: Longman.

Strevens, P. (1987). Interaction outside the classroom: using the community. In W. Rivers (Ed.) *Interactive language teaching* (pp. 170-176). Cambridge: Cambridge University Press.

尾崎明人・ネウストプニー, J.V. (1986).「インターアクションのための日本語教育：イマーションプログラムの試み」.『日本語教育』59号, 41-52.

ネウストプニー, J.V. (1981).「外国人場面の研究と日本語教育」.『日本語教育』45号, 30-40.

ネウストプニー, J.V. (1995).『新しい日本語教育のために』. 大修館書店.

1996年1月8日受
1996年6月20日改定版受
1996年11月1日改定2版受

Research Forum

Japanese EFL Learners' Test-Type Related Interlanguage Variability

Akihiro Ito
Hiroshima University

The purpose of this article is two-fold: 1) to investigate the effects of differences in test-types on the accuracy rates in interlanguage performance of Japanese EFL learners, and 2) to examine the reliability and validity of a grammaticality judgment test. Three grammar tests of relative clauses with three different test-types were assigned to 41 Japanese high school students. The tests were constructed on the basis of amount of attention to linguistic form from reviewing recent SLA works on task variation theory. From the results of the investigation, it is argued that 1) the participants showed different accuracy rates in the three test-types according to the expected order, and 2) the grammaticality judgment test showed relatively high reliability (rxx'=0.792) and moderate validity. The article also discusses the pedagogical implications of the findings and future direction of research.

この論文の目的は二つある。一つは日本人英語学習者の中間言語パフォーマンスへのテスト・タイプの効果を調査することであり、もう一つは文法性判断テストの信頼性と妥当性を検証することである。3種類の関係詞節に関するテストを41人の日本人高校生に実施した。テストは最近の第二言語習得研究におけるタスク要因理論を参考にして、言語形式への注意の度合を変数として作成された。調査の結果、3種類のテストにおいて仮説通りの順番で正確さの度合が異なり、文法性判断テストは比較的高い信頼性（rxx'=0.792)と中程度の妥当性を示した。本稿はさらにこの研究の教育への示唆と将来の研究の方向についても論じている。

When we try to estimate language learners' competence by measuring performance on a certain grammatical item, we often find it quite difficult to decide which test-type should be used. Recent research on second language acquisition theory has suggested that the accuracy rates of interlanguage performance systematically vary according to the kind of test-type[1] (Kameyama, 1987; Ohba, 1994a, 1994b). This phenomenon has been mainly explored in

an area of second language acquisition research called task variation theory (Ohba, 1987).

In the early to mid-1990s, researchers in the field of language testing claimed that research on language achievement tests had been neglected (Weir, 1993; Negishi, 1995) because most language testing professionals were more interested in measuring learners' general or overall proficiency.[2] Researchers tried to construct tests with potential for explaining learners' proficiency. As a consequence, little attention was being paid to test-type.[3] From the late 1970s through the 1980s, findings from SLA research implied that it was dangerous to measure learners' achievement level on a grammatical item or feature through use of a test containing only one test-type[4] (c.f. Kameyama, 1987; Nunan, 1992; Ohba, 1994a, 1994b). This was because of the reported systematic performance variability shown in foreign or second language learners' interlanguage according to test-type (Tarone, 1979, 1982, 1983, 1985, 1988; Sajjadi & Tahririan, 1992). Hence, the accuracy rate of the learners' interlanguage performance measurements can be affected by the test-type even if intended to measure a similar trait (Bachman, 1990; Ohba, 1994a, 1994b). This has served as the motivation for the present research. In this study relative clauses are the subject for three different test-types constructed to examine the effect of difference in test-type on Japanese EFL learners' interlanguage performance.

Test-type Classification

Recent research on test-types in SLA began with the dichotomous catergorization of test types into Natural Communication Tasks and Linguistic Manipulation Tasks (Dulay, Burt, & Krashen, 1982). A Natural Communication Task required learners to pay attention to content in order to use language for communication. A Linguistic Manipulation Task asked learners to pay attention to linguistic manipulation of form. This categorization was based on the *Monitor Model* (Krashen, 1977a, 1977b, 1978a, 1978b, 1981, 1982; Krashen & Terrell, 1983). The *Monitor Model* "predicts that the nature of second language performance errors will depend on whether monitoring is in operation" (Krashen, 1982, p. 152). Therefore, the Linguistic Manipulation Task, which encourages use of the monitor, measures to what degree learners have mastered grammatical rules and permits them to show higher accuracy. On the other hand, the Natural Communication Task, which does not permit monitor use, measures subconsciously-learned grammatical rules, which results in lower accuracy. In other words, Dulay, Burt, & Krashen (1982) suggest that test-types be classified rather like an on-off switch based on whether the test-type allows the monitor to be in operation or not.

Criticisms of test-type classification based on this dichotomy were quick to appear. Tarone (1983) first modified the dichotomy of test-types from the Labovian sociolinguistic perspective, arguing that there was no clear on-off point. She then proposed that the accuracy of interlanguage performance ranges on a continuum from vernacular to careful style based on the *Interlanguage Continuum Model*. Following this model, Hyltenstam (1983) proposed a more detailed categorization of experimental data, suggesting the following eight test-types:

a) Elicited production, often with pictorial stimuli, such as the *Berko Test*, the *Bilingual Syntax Measure*, guided composition;
b) Manipulation of given linguistic material such as sentence combining and sentence completion;
c) Intuition or grammaticality judgment test;
d) Introspection;
e) The cloze procedure;
f) Imitation;
g) Dictation or partial dictation; and
h) Translation. (p. 58)

Hyltenstam (1983), examining research on the relationship between interlanguage performance and selection of test-types, argued that it was impossible to divide test-types into two groups and reasonable to categorize them according to how much each required attention to linguistic form and content. According to this classification, as learners move toward translation, and pay greater attention to linguistic form, the resulting accuracy rate in the test is higher.[5]

Research on Interlanguage Variability

Most research on the effect of test-types on learners' interlanguage performance has not stemmed from research on testing but from other perspectives of SLA (Ohba, 1994a, 1994b). Ohba (1987, 1994a, 1994b) makes a rough division of what kind of grammatical items are chosen for the purpose of second language acquisition research, categorizing them into three types. The first, concerned with phonology, investigates the relationship between accuracy rate of a particular phoneme and the type of tasks provided to the participants (Dickerson, 1975; Beebe, 1980; Sato, 1985; Schmidt, 1987; Shizuka, 1993). The second examines the acquisition of morphology (Larsen-Freeman, 1975; Kameyama, 1987; Tomita, 1988; Inoi, 1991; Takamiyagi, 1991). The third discusses the

acquisition of syntax (Bailey, Eisenstein & Madden, 1976; Schmidt, 1980; Hueber, 1985). The results from these studies have supported the idea that the accuracy rate of language learners' performance changes according to the *Interlanguage Continuum Model,* though some morphological features such as articles have been shown not to follow it.

Recently, Ohba (1994b) investigated the effect of different test types on the interlanguage performance of Japanese EFL learners. The noteworthy points in this study were:

1. the use of a larger number of participants (N=370) in order to generalize the findings to a statistical population;
2. the division of participants by proficiency level (higher, average, lower) from scores on the *STEP* placement test (Obunsha, 1987, 1989) to determine whether higher-level learners show differences in accuracy rates;
3. the use of three test-types a) Grammaticality Judgment, b) Sentence Combining, and c) Picture Description;
4. the selection of a complex syntactic structure, relative clauses, to examine the effect of test-types because accuracy rates for morphology such as articles are easily influenced by discourse (Long & Sato, 1984; Tarone, 1985; Tarone & Parrish, 1988; Ohba, 1994a, 1994b), making it difficult to determine if the difference affects interlanguage performance.

Though Ohba (1994a, 1994b) made a significant contribution to this field, there were, nevertheless, a few drawbacks to his research. First, since different sentences were used for each test, it cannot be concluded that the scores were affected only by test-type. Second, in the Picture Description test, learners were asked to produce subject-type relativised sentences for sentences containing relative pronouns because these are considered easiest for L1 and L2 learners of English (Schachter, 1974; Keenan & Comrie, 1977; Gass, 1980) and avoided other kinds of relative clauses. The other test-types, Grammaticality Judgment test and Sentence Combining test, consisted of a mixture of sentences with four locations of the head noun phrase to be relativised: subject, direct object, object of preposition, and possessive. Therefore, though the Picture Description test was classified as not requiring attention to linguistic form, the characteristics of relative clauses may have resulted in higher than expected accuracy rates. Third, though Ohba said the low time pressure might have activated the learners'

monitor in the Picture Description test to explain the accuracy rate, Krashen (1985) said that time pressure may be unrelated to monitor activation. Therefore, it is difficult to view lack of time pressure as the reason for the increased accuracy rate.

In response to these concerns, the effects of test-types alone on learners' interlanguage performance need to be examined.

The Study

Purpose and Research Questions

The purpose of the present study was threefold: first, to replicate and expand Ohba's (1994a, 1994b) findings under more strictly controlled conditions; second, to examine the reliability coefficients of the three tests, and third, to determine what the grammatical judgment test employed by Ohba examines.

It was expected that the participants' interlanguage performance would vary according to Hyltenstam's (1983) theoretical framework.

Grammaticality judgment tests tend to be criticized (Ohba, 1994b), since some researchers (Ellis, 1991, 1994; Chaudron, 1983) are skeptical of their reliability. Moreover, some researchers question what grammaticality judgment tests measure in comparison to tests with different test-types. In spite of such criticism, grammaticality judgment tests are widely used because they are easy to construct and are believed to be a useful tool for eliciting learners' linguistic knowledge (Ohba, 1994b). To address these concerns, a grammaticality judgment test constructed on the basis of findings concerning relative clauses (Gass, 1980; Kawauchi, 1988) was used to examine the reliability. In order to determine what the grammaticality judgment test really examines, the correlations between the grammaticality judgment test and other tests, which target the same trait, but with different test-types, were calculated.

The research questions are:

1. Does the accuracy rate of the three tests follow the pattern of: Cloze > Grammaticality Judgment > Sentence Combining?
2. What does the Grammaticality Judgment measure? Is this test's reliability low?

The small sample ($N=41$) necessitated a conservative treatment of statistical analyses. Therefore, the alpha level for all statistical decisions was set at $a<0.01$.

Method

Subjects: The subjects in this study (*N*=41) were second year high school students (10th grade) enrolled in an English reading class at a high school attached to the Aichi University of Education. All were native speakers of Japanese. The average age was 16. All had completed at lease four years of formal English courses. The sample was thus homogeneous with regard to nationality, language background, educational level, and age. The group consisted of 21 males and 20 females. One male student was absent throughout this study. Though he later took the tests, his scores were not included.

In general, Japanese students are required to learn the usage of relative clauses: subject (SU), direct object (DO), indirect object (IO), and genitives (GEN), in junior high school, and relative clauses of preposition (OPrep) in high school. Therefore, it was concluded that the Ss had basic ability regarding use and comprehension of the English relative clause.

Materials: The following three 24-item relative clause tests were administered: 1) Cloze, 2) Grammaticality Judgment, and 3) Sentence Combining. These tests are a modification of the tests in Ohba (1994b) (see Appendix). To prevent use of only the easiest type of relative clause, subject type (SU), the same sentences were used for each test and the Picture Description test was replaced by the Cloze.[6]

In the process of constructing the three relative clause tests, careful attention was paid to the Noun Phrase Accessibility Hierarchy: SU (subject) > DO (direct object) > IO (indirect object) > OPrep (object of preposition) > GEN (possessive) > OComp (object of comparative particle) (Keenan & Comrie, 1977). The location of the head noun phrase is considered an influential component in the degree of difficulty associated with relative clauses in both L1 and L2 English acquisition (c.f. Schachter, 1974; Keenan & Comrie, 1977; Gass, 1980; Eckman, Bell, & Nelson, 1988; Akagawa, 1992; Sadighi, 1994; Aarts & Schils, 1995).

The tests: The 24 questions in the Sentence Combining test contained six pairs of sentences to be combined into sentences containing a relativised SU, six into sentences containing a relativised DO, six into sentences with a relativised OPrep, and six into sentences containing genitive cases. In the Cloze test, an appropriate relative pronoun, who, which, whose, whom, was required. In the Grammaticality Judgment test, a determination of each sentence's grammaticality, using either "O" or "X" as markers for correctness and incorrectness respectively, was required. For sentences judged incorrect, Ss were asked to make

necessary corrections.Typical errors in relative clauses are universal. They are categorized as: 1) relative clause marker omission, 2) pronoun retention, 3) wrong selection of relative clause marker, and 4) adjacency (Gass, 1980; Kawauchi, 1988). There were 12 correct and 12 incorrect sentences.

Test administration: In response to the shortcomings of recent studies, the ordering of the three test papers was taken into consideration. Since in the three tests the same sentences appear, the possibility of Ss memorizing the orthography to gain a higher score in later tests was considered. In order to reduce the potential order effect, the 24 items in each test were divided into three groups. Eight items from each of the tests were combined to make a 24-item test. In the first session, Cloze items 1-8, Grammaticality Judgment items 9-16, and Sentence Combining items 17-24 appeared; in the second session, Grammaticality Judgment items 1-8, Sentence Combining items 9-16, and Cloze items 17-24; and in the third session, Sentence Combining items 1-8, Cloze items 9-16, and Grammaticality Judgment items 17-24. Ss were allowed 20 minutes to complete each test. There was a one week interval between testing sessions, assumed to be long enough for Ss to forget some of the orthography and decrease the negative order effect. Ss were not told the dates of the sessions. Each session was conducted at the beginning of English reading classes by the instructor and his assistant. Though Ss were not informed of the purpose of the tests, they were encouraged to answer as many questions as possible. It is noteworthy that the Ss showed a great deal of interest on all the tests.

Scoring procedure: The tests were scored by the author. Scoring was based on whether the Ss had used the appropriate pronouns following an established criterion (Celce-Murcia & Larsen-Freeman, 1983; Quirk, Greenbaum, Leech, & Svartvik, 1985; Ohba, 1994a, 1994b). Therefore, local errors such as spelling mistakes were ignored as long as the meaning was clear.

Reliability estimation based on internal consistency: The Spearman-Brown split-half method was used to estimate the tests' reliability coefficients. The split-half method can be used when each test item is regarded as independent and also can contribute to the total score independently (local independence). In all three tests, each item is clearly independent. Thus, the split-half method is a permissible estimating procedure. The author scored the odd- and even-numbered items separately and first examined the Pearson's product-moment correlation (r) (Ito, 1996; Brown, personal communication, February 22, 1996). Each value was

then corrected for the reduction to half-test length using the Spearman-Brown prophecy formula ($rxx' = 2r/1+r$).

Results and Discussion

In this section, descriptive statistics of the tests are shown and the research questions addressed. Before discussing the results, however, two other aspects must be considered:

1. the effects of sample homogeneity on reliability and estimated correlations among the three tests, and
2. the small size of the sample (N=41).

Table 1: Reliability Coefficient, Mean, Maximum Score & Standard Deviation of Tests (N=41)

Tests	Reliability *rxx'*	Mean (*M*)	Max. Score	*SD*
CL	0.596	14.976	24	4.891
GJ	0.797	12.976	24	3.309
SC	0.693	10.732	24	6.490

CL=Cloze test; GJ=Grammaticality Judgment test; SC=Sentence Combining test.

Table 2: Analysis of Variance Summary Table (N=41)

Source	*SS*	*df*	*MS*	*F*-ratio	*p*
Subjects	2333.659	40	58.341		
Test-Type	498.260	2	249.124	24.215	<0.000
Error	823.073	80	10.288		
Total	3654.073	122			

Table 3: Multiple comparison test (Ryan's method) summary table (N=41)

Pair	*r*	Nominal Level	*t*	*p* level	Significance
CP-SC	3	0.003	6.920	0.000	s.
CP-GJ	2	0.007	2.823	0.005	s.
GJ-SC	2	0.007	4.097	0.000	s.

CL=Cloze test; GJ=Grammaticality Judgment test; SC=Sentence Combining test.

1. Does the accuracy rate of the three tests follow the pattern of: Cloze > Grammaticality Judgment > Sentence Combining?

Table 1 shows that in the relative clause tests the Cloze test had the highest mean ($M = 14.976$), the Grammaticality Judgment test a lower mean ($M = 12.976$), and the Sentence Combining test the lowest ($M = 10.732$). The accuracy rate seems to have changed according to the expected order. In order to determine statistical significance, one-way analysis of variance was performed. Table 2 shows an overall significant difference in the three test scores ($F(2,80) = 24.833$, $p < 0.000$). A multiple comparison test using Ryan's method was performed to determine where the difference lay. Table 3 shows it existed in each pair of the three tests at $p < 0.01$. Therefore, hypothesis 1 was confirmed.

2. What does the Grammaticality Judgment measure? Is this test's reliability low?

Table 1 shows the reliability coefficients of the three relative clause tests. Unexpectedly, the Grammaticality Judgment test showed the highest reliability among the three tests ($r = 0.798$), with reliability high enough for it to be regarded as a reliable testing device. The Cloze test ($r = 0.597$) and the Sentence Combining test ($r = 0.694$) showed moderate reliability. Ranked from highest to lowest in reliability coefficients, the reliability was: Grammaticality Judgment > Sentence Combining > Cloze.

Table 4: Correlations between each pair of three tests ($N = 41$)

Tests	r (exploratory rate %)	p
CP & GJ	0.626 (39.189)	<0.01
GJ & SC	0.696 (48.442)	<0.01
SC & CP	0.698 (48.720)	<0.01

CL = Cloze test; GJ = Grammaticality Judgment test; SC = Sentence Combining test.

Results indicate that a Grammaticality Judgment test can be relatively reliable. However, validity also needs to be investigated. Table 4 displays the correlation for each pair of three tests. The correlations measured show almost the same magnitudes: $0.60 < r < 0.70$. Results reveal that each pair of tests shared the same amount of trait or ability needed for producing or comprehending relative clauses (exploratory rate or coefficients of determination: $r^2 = 39.189$ to 48.720%). [The square of

value of r (r^2) •100 indicates what percentage of similar traits or abilities each pair of tests share]. The Grammaticality Judgment test shows a relatively high reliability coefficient and moderate correlation with the other two test-types.

Though Grammaticality Judgment tests are still controversial with regard to reliability and validity, this research indicates they have far-reaching potential as reliable and valid elicitation tools. However, test designers must be aware of the universal error types or typical errors from learners' L1 transfers in order to construct appropriate tests. In addition, Table 1 reveals that the Grammaticality Judgment test shows the lowest standard deviation here, implying it has limited discriminative ability.

Conclusions and Pedagogical Implications

The results of this study indicate that, for Japanese EFL high school students: 1) the accuracy rate follows the expected pattern (Cloze > Grammaticality Judgment > Sentence Combining); and 2) unexpectedly, the Grammaticality Judgment showed fairly high reliability, with moderate correlations between the two other test-types. However, since its discriminative ability seems limited, it should be used with extreme care.

In a pedagogical sense, the results indicate that the manifestation of learners' interlanguage competence, their performance, varies according to test-type. As a consequence, teachers may well underestimate or overestimate learners' knowledge or ability to use a grammatical item if they rely on only one test-type. Moreover, in order to characterize the learners' actual abilities in the target language, it is necessary to employ a variety of test-types.

Limitation and Suggestions for Further Research

It should be acknowledged that one of the limitations of the present investigation is that it focused only on the Ss' performance. Thus, the results can be generalized only for Japanese students. However, in many studies conducted in the past, various language backgrounds, ages, and educational backgrounds were mixed. As a result, the findings have often been hard to interpret because they can only be generalized to the single situation in which the data was collected. In addition, the results of this study may be influenced by two internal and external factors:

1. the nature of reliability in measures in general, and
2. restrictions in the range of ability that was sampled in the investigation.

Generally, tests are not simply reliable and valid but they can be reliable and valid for specific types of students and specific ranges of ability (Brown, personal communication, February 22, 1996). In this regard this research should be replicated with a larger sample of participants from a much wider population.

The following three general research questions are posed in the hope that other researchers will pursue further investigations.

1. Does the proficiency level affect the magnitude of inter-language variability with regard to accuracy rates? If so, does the degree of the variability in the target language decrease as proficiency level increases (i.e. higher level learners < average level learners < lower level learners?).[7]
2. What really causes the difference in accuracy rates? The amount of attention to linguistic form by monitoring? Or the difference in cognitive processes and demands required of subjects?
3. How high is the construct validity of each test? That is, does each of the tests measure what it is constructed to measure?

Acknowledgments

The author sincerely thanks Norihisa Matsumoto, Shogo Miura, Yoshito Nakamura, Tadashi Nishida, Nobukazu Matsu-ura, Hiroshima University, and J.D. Brown, University of Hawaii, for their valuable critical comments on earlier drafts of this paper. Thanks also go to Hiromasa Ohba, Health Sciences University of Hokkaido; Kenji Ohtomo, University of Tsukuba; and Susan Gass, Michigan State University for their communications on this topic, as well as Toshifumi Takizawa, Takao Imai, Yuya Koga, Shizuya Tara, Hiroyuki Araki, Hidetoshi Inoue, and Fumiyo Yoshikawa, for their encouragement and cooperation, and two anonymous reviewers of the JALT Journal *for their valuable comments. I alone, however, am responsible for the analysis herein.*

Akihiro Ito, M.A., Hiroshima University, is currently a Ph.D. candidate in the Department of English Language Education, Hiroshima University. His research interests include language testing, evaluation, and second language acquisition research methodology.

Notes

1. Throughout this paper, "test-type" refers to any task type. Tasks for data elicitation are widely used in language testing conditions. Terms such as test format, test method, or method facet can refer to the concept of test-type used here since there has been a great amount of variability in classification.
2. Following Bachman's (1990) definition, "performance" is used as an indicator of a long-standing ability or competence, a person's knowledge of the language, which can be estimated indirectly by a test score and its valid interpretation (p. 33). On the basis of Bachman's definition, I take the position that competence is basically homogeneous and unitary, and that interlanguage variability manifested in different test-types is essentially a phenomenon of performance. The variability of performance can be observed by examination of change in accuracy rate or score on tests.
3. Some language testing and/or curriculum experts (Brindley, 1989) claim that the distinction between proficiency and achievement is not clear-cut. However, in my view, an achievement test is given to language learners at the end of a program to check if they have mastered the targeted items or skills. In this sense, an achievement test is, as Alderson, Clapham, and Wall (1995, pp. 11-12) argue, similar to a progress test. Proficiency tests are not based on language programs or classes. The main purpose of proficiency tests is to examine overall language ability.
4. For overviews of research on task variation theory, see Ohba (1987, 1994b).
5. Tarone's (1983) categorization of elicitation tools is comprised of tasks which are more differentiated for Hyltenstam (1983). Other test-types such as cloze tests or translation test-types in Hyltenstam's classification of experimental data can be located on a continuum with regard to accuracy rate, though some must be placed in the careful-style area.
6. Some might question if the cloze test-type in Hyltenstam (1983) allows higher accuracy than grammaticality judgment tests or sentence combining tests. DeKeyser (1990), however, argues that the fill-in-the-blank test-type, which is similar to the cloze, works as a more valid measure of monitored knowledge. He discusses the effectiveness of the fill-in-the-blank test-type for examining monitored knowledge compared to the other test-types. In this regard, my interpretation of Hyltenstam's theoretical framework is a combination of Tarone's (1983) interlangauge continuum model and DeKeyser's proposals.
7. Reexamination of the data collected to determine the effects of proficiency level on the magnitude of variability of accuracy rates in participants' interlanguage performance is currently underway.

References

Aarts, F., & Schils, E. (1995). Relative clauses, the accessibility hierarchy and the contrastive analysis hypothesis. *International Review of Applied Linguistics, 33*, 47-63.

Akagawa, Y. (1992). Avoidance of relative clause by Japanese high school students. *JACET Bulletin, 23*, 1-18.

Alderson, J.C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.

Bachman, L.F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bailey, N., Eisenstein, M., & Madden, C. (1976). The development of wh-questions in Adult second language learners. In Faneslow, J., & Crymes, R. (Eds.), *On TESOL '76* (pp. 1-9). Washington, DC: TESOL.

Beebe, L.M. (1980). Sociolinguistic variation and style shifting in second language acquisition. *Language Learning, 30*, 178-194.

Brindley, G. (1989). *Assessing achievement in the learner-centered curriculum*. Sydney: National Centre for English Language Teaching and Research.

Celce-Murcia, M., & Larsen-Freeman, D.E. (1983). *The grammar book: An ESL/EFL teacher's course*. Rowley, MA: Newbury House.

Chaudron, C. (1983). Research on metalinguistic judgments: A review of theory, methods, and results. *Language Learning, 33*, 343-377.

DeKeyser, R. (1990). Towards a valid measurement of monitored knowledge. *Language Testing, 7*, 147-157.

Dickerson, L.J. (1975). The learner's interlanguage as a system of variable rules. *TESOL Quarterly, 9*, 401-407.

Dulay, H., Burt, M., & Krashen, S.D. (1982). *Language two*. Oxford: Oxford University Press.

Eckman, F.R., Bell, L., & Nelson, D. (1988). On the generalization of relative clause instruction in the acquisition of English as a second language. *Applied Linguistics, 9*, 1-20.

Ellis, R. (1991). Grammar judgments and second language acquisition. *Studies in Second Language Acquisition, 13*, 161-186.

Ellis, R. (1994). Data, theory, and applications in second language acquisition research. In Ellis, R. (Ed.), *The study of second language acquisition* (pp. 669-691). Oxford: Oxford University Press.

Gass, S. (1980). An investigation of systematic transfer in adult second language learners. In Scarcella, R.C., & Krashen, S.D. (Eds.), *Research in second language acquisition* (pp. 132-141). Rowley, MA: Newbury House.

Hueber, T. (1985). System and variability in interlanguage syntax. *Language Learning, 35*, 141-163.

Hyltenstam, K. (1983). Data type and second language variability. In Rongbon, H. (Ed.), *Psycholinguistics and foreign language learning* (pp. 57-74). Abo (Turku), Finland: Abo Akademi.

Inoi, S. (1991). Variation in interlanguage with special reference to articles and pronouns. *Annual Review of English Language Education in Japan, 2*, 1-10.

Ito, A. (1996). Testing English tests: A language proficiency perspective. *JALT Journal, 18*, 183-197.

Ito, A. [in press]. A study on the variability of test performance of Japanese EFL learners: A combination of two theoretical frameworks. *CELES Bulletin, 26*.

Kameyama, T. (1987). *Bunpo tesuto niokeru tesuto keishiki niyoru keitaiso seitoritsu no henka* (The change of accuracy rate in grammar tests by the

difference of test-types). *CELES Bulletin, 17*, 248-252.

Kawauchi, C. (1988). Universal processing of relative clauses by adult learners of English. *JACET Bulletin, 19*, 19-36.

Keenan, E.L., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry, 8*, 63-99.

Krashen, S.D. (1977a). The monitor model for adult second language performance. In Burt, M., Dulay, H. & Finocchiaro, M. (Eds.), *Viewpoints on English as a second language* (pp. 152-161). New York: Regents.

Krashen, S.D. (1977b). Some issues relating to the monitor model. In Brown, D.H., Yorio, C.A., & Crymes, R.H. (Eds.), *On TESOL '77* (pp. 144-158). Washington, DC: TESOL.

Krashen, S.D. (1978a). Individual variation in the use of the monitor. In Ritchie, W.C. (Ed.), *Second language acquisition research: Issues and implications* (pp. 175-183). New York: Academic Press.

Krashen, S.D. (1978b). The monitor model for second language acquisition. In Gingras, R.C. (Ed.). *Second language acquisition and foreign language teaching* (pp. 1-26). New York: Center for Applied Linguistics.

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

Krashen, S.D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon Press.

Krashen, S.D. (1985). *The input hypothesis: Issues and implications*. London: Longman.

Krashen, S.D. & Terrell, T. (1983). *The natural approach: Language acquisition in the classroom*. Oxford: Pergamon Press.

Larsen-Freeman, D.E. (1975). The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly, 9*, 409-419.

Long, M.H., & Sato, C.J. (1984). Methodological issues in interlanguage studies: An interactionist perspective. In Davies, A., Criper, C., & Howatt, A.P.R. (Eds.), *Interlanguage* (pp. 253-279). Edinburgh: Edinburgh University Press.

Negishi, M. (1995). *Ri-dingu no tesuto to hyoka* (Reading test and evaluation). *Eigokyoiku (The English Teachers' Magazine), 1*, 29-31.

Nunan, D. (1992). *Research methods in language learning*. Cambridge: Cambridge University Press.

Ni kyu yosou mondaishu (Collections of Second Grade STEP Tests). (1987). Tokyo: Obunsha.

Ni kyu yosou mondaishu (Collections of Second Grade STEP Tests). (1989). Tokyo: Obunsha

Ohba, H. (1987). *Tasuku barieshun riron ni tsuite* (On the task variation theory). *CELES Bulletin, 17*, 43-49.

Ohba, H. (1994a). *Nihonjin eigo gakushusha no chukangengo kahensei: Tasuku keishiki no kanten kara* (Japanese EFL learners' interlanguage variability: With reference to task-type). *CELES Bulletin, 24*, 187-192.

Ohba, H. (1994b). *Tesuto keishiki no chigai niyoru eigo gakushusha no pafomansu no kahensei* (Task-related variability in interlanguage by Japanese EFL learn-

ers). *STEP Bulletin, 6*, 34-48.

Powers, D.E. (1982). Selecting samples for testing the hypothesis of divisible versus unitary competence in language proficiency. *Language Learning, 32*, 331-335.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. London: Longman.

Sadighi, F. (1994). The acquisition of English relative clauses by Chinese, Japanese, and Korean adult native speakers. *International Review of Applied Linguistics, 32*, 141-153.

Sajjadi, S. & Tahririan, M.H. (1992). Task variability and interlanguage use. *International Review of Applied Linguistics, 30*, 35-44.

Sato, C.J. (1985). Task variation in interlanguage phonology. In Gass, S., & Madden, C.G. (Eds.), *Input on second language acquisition* (pp. 181-196). Rowley, MA: Newbury House.

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

Schmidt, M. (1980). Coordinate structures and language universals in interlanguage. *Language Learning, 30*, 397-416.

Schmidt, R.W. (1987). Sociolinguistic variation and language transfer in phonology. In Ioup, G., & Weinberger, S.H. (Eds.). *Interlanguage phonology* (pp. 365-377). Rowley, MA: Newbury House.

Shizuka, T. (1993). Task variation and accuracy predictors in interlanguage phonology production. *Bulletin of Kanto-Koshin-etsu English Language Education Society, 7*, 63-77.

Takamiyagi, T. (1991). A study of task related variation in interlanguage by Japanese university students in an instruction only environment. *Joetsu University of Education: Kenkyu Ronshu (Bulletin of Language Studies), 6*, 47-64.

Tarone, E. (1979). Interlanguage as chameleon. *Language Learning, 29*, 181-191.

Tarone, E. (1982). Systematicity and attention in interlanguage. *Language Learning, 32*, 69-84.

Tarone, E. (1983). On the variability of interlanguage systems. *Applied Linguistics, 4*, 142-164.

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 Publishers.

Tarone, E. & Parrish, B. (1988). Task-related variation in interlanguage: The case of articles. *Language Learning, 38*, 21-44.

Tomita, Y. (1988). *Nihonjin kokosei no chukan gengo nitsuite no ichikousatsu: Keitaiso shutoku junjo kenkyu* (A study on the high school students' interlanguage: With reference to the acquisition order of morphologies). *CELES Bulletin, 18*, 208-213.

Weir, C.J. (1993). *Understanding and developing language tests*. London: Prentice-Hall International.

(Received Jan. 3, 1996; revised Oct. 7, 1996)

Appendix: Test sentences

Cloze

1. The policeman has caught the girl ( ) stole the car.
2. The author ( ) books I haven't read yet is well known.
3. The magazine from ( ) I got the information is Newsweek.
4. He still had the pen ( ) I had given to him.
5. I came across some students ( ) names I couldn't remember.
6. She paid the man from ( ) she had borrowed the money.
7. The wine ( ) you brought to our party was excellent.
8. You will receive the kind letter ( ) your mother wrote.
9. The book to ( ) I referred can be obtained from the library.
10. The girl ( ) is sitting at the reception desk is pretty.
11. The city from ( ) that boy came is far from here.
12. I met a friend ( ) mother was a famous designer.
13. The building ( ) stands near the lake is our hotel.
14. The man ( ) feet were very large has just bought new shoes.
15. I need someone ( ) can help me clean the house.
16. The woman ( ) John will marry next month is Japanese.
17. The man ( ) was injured in the accident is in the hospital.
18. I saw a man ( ) bag was the same as mine.
19. I have just found the key ( ) I lost yesterday.
20. The book ( ) I bought the other day is interesting.
21. The woman with ( ) he was talking was Mrs. Miller.
22. The police interviewed the lady from ( ) the diamonds had been stole.
23. The boy ( ) essay I corrected has entered Hokkaido University.
24. I know the children ( ) are playing in the yard.

Grammaticality Judgment

1. ( ) The policeman has caught the girl stole the car.
2. ( ) The author whose books I haven't read yet is well known.
3. ( ) The magazine which I got the information from it is Newsweek.
4. ( ) He still had the pen which I had given to him.
5. ( ) I came across some students names I couldn't remember.
6. ( ) She paid the man from whom she had borrowed the money.
7. ( ) The wine was excellent which you brought to our party.
8. ( ) You will receive the kind letter whose your mother wrote.
9. ( ) The book to which I referred can be obtained from the library.
10. ( ) The girl who is sitting at the reception desk is pretty.
11. ( ) The city is far from here which that boy came from.
12. ( ) I met a friend which mother was famous designer.
13. ( ) The building which stands near the lake is our hotel.
14. ( ) The man has just bought shoes whose feet were very large.

15. ( ) I need someone who can help me clean the house.
16. ( ) The woman whom John will marry next month is Japanese.
17. ( ) The man is in the hospital who was injured in the accident.
18. ( ) I saw a man whose bag was the same as mine.
19. ( ) I have just found the key which I lost yesterday.
20. ( ) The book which I bought it the other day is interesting.
21. ( ) The woman with whom he was talking was Mrs. Miller.
22. ( ) The police interviewed the lady from which the diamonds has been stolen.
23. ( ) The boy whose essay I corrected has entered Hokkaido University.
24. ( ) I know the children which are playing in the yard.

Sentence Combining

1. The policeman has caught the girl. She stole the car.
2. The author is well known, I haven't read his books yet.
3. The magazine is Newsweek. I got the information from it.
4. He still had the pen. I had given it to him.
5. I came across some students. I couldn't remember their names.
6. She paid the man. She had borrowed the money from him.
7. The wine was excellent. You brought it to our party.
8. You will receive the kind letter. Your mother wrote it.
9. The book can be obtained from the library. I referred to it.
10. The girl is pretty. She is sitting at the reception desk.
11. The city is far from here. The boy came from it.
12. I met a friend. His mother was a famous designer.
13. The building is our hotel. It stands near the lake.
14. The man has just bought new shoes. His feet were very large.
15. I need someone. Someone can help me clean the house.
16. The woman is Japanese. John will marry her next month.
17. The man is in hospital. He was injured in the accident.
18. I saw a man. His bag was the same as mine.
19. I have just found the key. I lost it yesterday.
20. The book is interesting. I bought it the other day.
21. The woman was Mrs. Miller. He was talking with her.
22. The police interviewed the lady. The diamonds had been stolen from her.
23. The boy had entered Hokkaido University. I corrected his essay.
24. I know the children. They were playing in the yard.

Codeswitching in EFL Learner Discourse

Ethel Ogane
Tokyo YMCA

L1 use by L2 learners has been couched in negative terms like "resort to" and "fall back on." However, L1 use can be looked at in a more positive way. The alternative use of the learners L1 and their approximation of the TL may be termed codeswitching (CS). This study looks at CS in five female Japanese EFL learners. Analysis of spoken data and insights from subjects' verbal reports suggest that CS is discourse related. CS helps learners manage and smooth the flow of conversation, and allows them to express their dual identities of student and individual in the classroom.

第二言語学習者による第一言語の使用は「第一言語に頼る」という表現にも見られるように否定的に捉えられてきた。しかしながら第一言語の使用はもっと肯定的に捉えることも可能である。第一言語と目標言語の近似値である中間言語との交替はコード切り替えと考えることもできる。本研究は、5人の日本人女性英語学習者によるコード切り替えを検討した。被験者の会話データと内省報告を分析したところ、コード切り替えは談話に関連していることが示唆された。コード切り替えは学習者が会話の流れを管理し円滑にする助けとなり、教室における学生と個人という二重のアイデンティティを表現する手段となっている。

Teachers may oppose the English Only movement on a sociopolitical level but accept English-only classroom management practices as common and natural (Auerbach, 1993). L1 use is couched in negative terms like "resort to" and "fall back on." In the literature on communication strategies, L1 use is labeled as a "compensatory strategy," problem solving in nature, used to overcome "insufficient linguistic resources" (Faerch & Kasper, 1983, p. 46). Tarone, Cohen and Dumas (1983) label it an "avoidance strategy" (p. 11). This attitude toward L1 use, as evidenced by the words "compensatory" and "avoidance," seems to focus negatively on learner interlanguage as linguistically inferior to the target language. Kasper and Blum-Kulka (1993) suggest that "instead of viewing interlanguage pragmatic knowledge and behavior as deficient in terms of native norms, we need to consider its functionality and inner justification" (p. 160).

L1 use by learners may be described as codeswitching (CS) behavior. That is learners switching back and forth between their L1 and their approximation of the L2. Legenhausen (1991) suggests that CS in learners be compared and contrasted with that of more proficient bilingual speakers. Codeswitching, "the alternative use by bilinguals of two or more languages in the same conversation" (Milroy & Muysken, 1995, p. 7), was thought to be abnormal, random behavior and indicative of a lack of language proficiency. The last 25 years of research on CS has contributed to its "rehabilitation" (Gardner-Chloros, 1991, p. 59). This research has demonstrated that language alternation is not arbitrary behavior but a type of "skilled performance" (Myers-Scotton, 1993c, p. 47), an important and "extremely common characteristic of bilingual speech" (Grosjean, 1982, p. 146).

Types of CS

The types of switching identified in the literature include: *tag*, *intersentential*, and *intrasentential* (Appel & Muysken, 1987; Poplack, 1980; Romaine, 1995). Tag-switching is the placing of a tag, an exclamation, a formulaic expression, or a discourse particle from one language into an utterance which is, except for that item, entirely in another language. Tag-switching has been referred to as emblematic switching because the switched item serves to identify an otherwise monolingual utterance as bilingual in character. Examples[1] of tag-switches are:

1) *Vendia arroz* (He sold rice) 'n shit. (Poplack, 1980, p. 589)
2) I guess it's good *yo ne* (right)? (Nishimura, 1995, p. 169)

An intersentential switch occurs at clause or sentence boundaries in one of the languages or between speaker turns (Romaine, 1995), as in:

3) Sometimes I'll start a sentence in English *y termino en espanol* (and finish it in Spanish). (Poplack, 1980, p. 594)

Intrasentential switching, sometimes called code mixing, involves a switch within a clause or sentence and sometimes even within word boundaries. Examples are:

4) Why make Carol *sentarse atras pa'que* (sit in the back so) everybody has to move *pa'que se salga* (for her to get out)? (Poplack, 1980, p. 589)
5) And it's hard, 'cause me- *nanka, moo, hon o nanka yomu to,* cover-to-cover *yomanakattara*, if I stop *dokka de,* I forget the story. (And it's hard, because a person like me, when I read a

book, unless I read it cover-to-cover, if I stop at some point, I forget the story.) (Nishimura, 1995, p. 178)

CS Research

The approach to CS has been interdisciplinary (Appel & Muysken, 1987; Milroy & Muysken, 1995; Tabouret-Keller, 1995). Researchers have attempted to locate and explain constraints on CS in terms of: a) how the bilingual mind works (psycholinguistic); b) the formal properties of linguistic systems (structural); and c) the social, historical and interactive processes of individuals and groups in language contact situations (sociolinguistic). Psycholinguistic aspects are not reviewed here.[2]

Structural aspects: Early studies tried to explain why particular points were chosen for switches. The free-morpheme constraint, proposed by Sankoff and Poplack (as cited in Romaine, 1995) predicted that "a switch may not occur between a bound morpheme and a lexical form unless the lexical form has been phonologically integrated into the language of the morpheme" (Romaine, 1995, p. 126). The equivalence constraint (Lipski, 1978; Poplack, 1980) held that a switch may occur where the juxtaposition of forms from two codes does not violate a syntactic rule of either code.

Recent work on CS has taken different approaches. The Matrix Language-Frame (MLF) (Myers-Scotton, 1993a) claimed production-based switching with constraints occurs not at the surface phrase structure but at an abstract level, the "mental lexicon" of the bilingual (p. 485). Romaine (1995), suggesting that "a mixed code has its own rules and constraints," said a switched item may not be "predictable from the individual constituent structure rules of the two systems in contact" (p. 160). Researchers also asked whether constraints were language-specific or universal (Romaine, 1995).

The study of linguistic constraints involves differentiating between borrowing and codeswitching. Grosjean (1982) defined borrowing as "a word or short expression that is adapted phonologically and morphologically to the language being spoken" (p. 308). However, problems of overlap in the actual diagnosis of whether an individual case is a switch or loan occurred, as when a speaker pronounces all words, borrowed or otherwise, in the same accent. Poplack (1988) maintained there is a dichotomy which could be worked out methodologically. Gardner-Chloros (1995) argued that CS is not "separable, either ideologically or in practice from borrowing, interference or pidginisation" (p. 86). Myers-Scotton (1993b) urged that codeswitching and borrowing be seen as part of a continuum.

Sociolinguistic aspects: CS was thought to be "part of the performance of the imperfect bilingual, motivated by inability to carry on a conversation" (Myers-Scotton, 1993c, p. 47) and not a serious research topic. A study of Norwegian dialects (Blom & Gumperz, 1972) presented CS as a legitimate topic of research and stimulated research on CS between languages (Myers-Scotton, 1993c). Two general kinds of CS were posited (Blom & Gumperz, 1972; Gumperz, 1982). In *situational* switching, speakers switch when they perceive or when there is a change in setting, topic of conversation, or participants. In *conversational* or *metaphorical* switching, speakers create meaning by switching. Gumperz (1982) suggested that CS is a discourse strategy used by bilinguals in much the same way monolinguals use style shifting and prosody. In addition, Gumperz (1982) distinguished between a "we code" and a "they code," usually the minority and majority languages respectively. Use of the "we code," implying intimacy and informality, is associated with in-group and personal activities, while the "they code," symbolizing authority and objectivity, marks more formal, less personal, out-group relations. Other CS functions included quotation, addressee specification, interjections, reiteration, and message qualification (Gumperz, 1982).

Other researchers have studied the functional aspects of CS. Poplack (1980) maintained that specific instances of switching in the smooth, rapid Spanish/English CS by Puerto Ricans in New York City cannot be assigned discourse functions. Rather, CS was the norm for this bilingual community. Poplack (1988) found a different pattern of CS among French/English bilinguals in the Ottawa-Hull community in Canada. These speakers drew attention to their CS by "repetition, hesitation, intonational highlighting" and "metalinguistic commentary" (p. 230). Language switching marked specific functions, including switching to provide an apt expression—*mot juste* switching (p. 228); to fill a lexical gap; to bracket or call attention to an English switch; and to explain, specify or translate.

The study of in-group speech repertoire of second generation Japanese-Canadian (Nisei) bilinguals (Nishimura, 1995) identified three speech varieties. In the basically Japanese variety, used with native Japanese speakers, CS serves to fill lexical gaps. In the basically English variety, used with each other, CS allows the speakers to express their shared ethnic identity. The mixed variety was used when the interlocutors included both native Japanese and Nisei. Nishimura (1995) categorized CS in the mixed variety into four groups: a) functions related to interactional (speaker/hearer) processes where Niseis tried to "reach

out" (p. 167) to both types of listeners or to enhance rapport; b) discourse organization functions where CS marked the beginning or ending of a frame or introduced a topic; c) functions carrying stylistic effects where Niseis marked quotations with a switch in order to make their speech "more lively" (p. 177); and d) functionally neutral CS where the motivation was not clear.

Some researchers have attempted to move on from a descriptive approach listing CS functions to a prescriptive approach treating CS patterns within an explanatory framework or model. Scotton and Ury (1977), who looked at Swahili/Luyia/English CS, explained CS as an extension of a speaker in terms of the relationships between participants and subject and the social meaning of a language choice. By using a switch, the speaker may redefine the social arena, or set of norms, of an interaction. By continually switching, the speaker may avoid specifying the social meaning of the interaction. In her study of French/Alsatian CS, Gardner-Chloros (1991) showed how switching was used by colleagues in work situations to create solidarity in spite of gaps in Alsatian competence, especially for younger speakers, and the lack of appropriate technical terms in Alsatian. CS "is connected with individual factors which concern people's linguistic histories as well as their personalities" (Gardner-Chloros, 1991, p. 184).

CS Research in the Classroom

Early research into classroom CS concentrated on the "communicative functions of code-switching in teacher-led talk and on the frequency with which particular languages were employed to perform different functions" (Martin-Jones, 1995, p. 90-1). Later studies looked at the "sequential flow of classroom discourse" and the interactional work of teachers and students from a conversational analytic approach with an ethnographic grounding (Martin-Jones, 1995, p. 91). This included studies on CS in EFL and ESL classrooms. Martin-Jones (1995) cited two studies by Lin (1988, 1990) on CS in Anglo-Chinese secondary schools in Hong Kong. In her 1988 classroom study, Lin found that the teacher, a Cantonese/English bilingual, used CS "as a communicative resource to signal unspoken social meanings" (p. 84). Fotos (1994, 1995) looked at CS in EFL classrooms in Japan. In her analysis of Japanese/English CS by Japanese university EFL students, Fotos (1995) found that the learners used switching into Japanese for emphasis, repetition, or clarification, as well as to signal that a mistake had been made in English and a repair was to follow. These functions call attention to the speakers' English utterances.

Foreign and second language learners may be thought of as incipient bilinguals (Kasper, 1994) and EFL/ESL classrooms as bi/multilingual communities. This study analyzes CS in a group of Japanese adult EFL learners to further understand the use of CS. What functions might L1 (Japanese) use or CS serve for developing bilingual speakers?

The Study

Method

Subjects: Five female students, Maki, Remi, Emi, Fumi, and Kono (not their real names), comprising an elementary level "conversation" class at a small private language school voluntarily participated. All had studied English at the secondary level and two had had a year of post-secondary lessons. They had attended language classes for three to seven years. Their ages ranged from 45 to 69.

Data Collection: Data were collected from the subjects' class sessions, following Gardner-Chloros' (1991) suggestion that naturally occurring classsroom discourse rather than language elicited by experimental tasks is better suited for the study of CS. Data from retrospective oral interviews were also collected. Retrospective data were included for two reasons: first, this researcher hoped that the individual interviews would help raise the subjects' awareness of L1 use in the classroom and stimulate reflection, and second, that the verbal reports, "the learners' reports of their own intuitions and insight," would serve to "complement" (Cohen, 1987, p. 82) the classroom discourse data.

Class sessions: Fourteen two-hour class sessions were video and audio taped from May to October, 1994. This duration was selected to ensure that Ss would become accustomed to being recorded. Three sessions, from June and September, were transcribed. Only discourse involving the entire class was transcribed; pair work was not included. During pair work, the Ss inevitably spoke Japanese, but it was impossible to transcribe the separate pair conversations.

Transcription of the data is in standard English orthography. The Japanese switches, in italics, translated by the author[3] into idiomatic English, appear in parentheses. Significant contextual information appears in brackets.

Interviews: Each subject was individually interviewed twice. All interviews were audio-taped. Each interview took place directly after a lesson and

lasted from one to two hours. Biographical data were collected in the first interview and retrospective data in the second.

During the first interview, Ss were asked about their English educational background and reasons for studying. Ss were told that they should feel free to use either Japanese or English and to request clarification whenever they did not understand. The researcher spoke mostly in English. During the interview, subjects were asked for comments on their use of Japanese in class.

During the second interview, the video tape of the lesson that had just taken place was shown. Subjects were asked to listen for instances when they used Japanese and to try to recall why they switched and what language they were thinking in when the switch occurred. Subjects were told that they could push the pause button whenever they wanted to make an observation or remembered something. I also tried to elicit comments from them on their CS.

Results and Analysis

Switches from one speaker's utterance to another's, and those within one speaker's utterance, were counted. Since the focus of this study is use of the L1, only switches from English to Japanese were considered.[4]

Table 1: Analysis of Switches to Japanese

Type of Switch	Frequency	%
Tags	39	8%
Interjections/Short-fixed Expressions	43	9%
wa/ga	3	1%
janakute	9	2%
Conjunctions	13	3%
ja	17	4%
Adjectives/Adverbs	18	4%
Nouns	45	10%
Fillers	48	10%
nan to iu	61	13%
Phrases/Sentences	164	35%
ni naru/ni suru	12	3%
Total	472	102%*

*Note: The Total equals more than 100% due to rounding.

Switches into the L1 designating place names and proper nouns were not included. Also excluded were loan-word status items like "bazaar" and "seminar" when they were used as attempts or approximations toward English. A phonological basis for determining whether a word was English or Japanese was not used since much of the subjects' interlanguage is highly accented.

As Table 1 shows, the learners frequently produced tag, interjection, noun, filler, and *nan to iu* (how do you say . . .) switches. Phrases and sentences were also a favored switch. Less fluent or L1 dominant bilinguals have been shown to prefer tag and single-item switches (McClure, 1981; Poplack, 1980). Fotos (1990), in her study of Japanese-English bilingual children's CS, found that single item and sentence switches were a "significant number of the total, 107 switches out of 153" (p. 84).

Tags: The most frequently used tag was *ne*. More than half of the tag switches were made by one learner, Fumi, who may have used *ne* as a mini-confirmation check (example 1) and as a repair acceptance marker (example 2). She also may have used *ne* to gain thinking time or keep her turn, saying she tried to use English as much as possible but felt that her English did not come out quickly enough.

1. Fumi: Matterhorn. But eh I can uh I can get to the, I could *ne*, I could get to the uh *eh* town . . .
2. Fumi: Healthness, healthy, mmm uh I I don't uh got a ill. I don't got after? uh after I
   Remi: After that.
   Fumi: After that, *ne*. After that I don't got the ill.

Subjects also used the tags *deshoo*, *kana*, and *ka*.

Interjections and short-fixed expressions: Switched interjections included *masaka*! (no kidding!), *naruhodo* (I see), and *e*? (what?); expressions included *gomen* (sorry) and *doomo arigatoo* (thank you). Other examples are:

3. T: I wore the earrings that I bought at the bazaar yesterday.
   Maki: *Maa!* (Oh!) [laughs] Many earrings you bought.
4. Fumi: One, two, three, four, five, six, seven, eight *ne*, plus eight.
   Kono: mmm
   Maki: *Sugoi!* (Wow!)

Maki and Remi reported in interviews that they were very aware of their Japanese classmates and sometimes used Japanese without thinking. Remi thought that she sometimes switched because she felt she had to respond in Japanese to her peers. Maki said she thought she used Japanese to be considerate of her classmates. The students appear to be using these switches to express personal feelings and to show solidarity with their peers. It is interesting to note, however, that in Example 3 Maki used an interjection switch in response to the teacher.

Wa *and* ga: There were three instances of *wa/ga* switches in the data. *Ga* is the subject marker and *wa* is the topic marker in Japanese.

5. Maki: Finish *ga*? (How about the finish?)
6. Emi: Finish *ga* October 14. (We finish on October 14.)
7. Maki: *Anoo* my lend *wa* very long time. Thank you very much. (I have borrowed this a very long time.)

Fotos (1994) also found *wa/ga* switches in her college student CS data and cites them as evidence against the free-morpheme constraint because the markers should be connected to Japanese topics and subjects. However, Fotos (1990) noted in her earlier children's CS study that these Japanese particles may not be bound "in the sense that bound morphemes exist in English or Spanish" (p. 88).

Nishimura (1989) found that second generation Japanese-American and Japanese-Canadian bilingual speakers, who all said they were better in English than in Japanese, stated a topic in Japanese then switched to a comment in English. The speakers' "thinking process is topic-comment" with the topic part realized in Japanese and the comment produced in English (p. 376). In Nishimura's (1995) Japanese-Canadian Nisei study, bilinguals speaking in the mixed variety marked topics with wa, using it to imply a contrast in order to "to single out something in a discourse" and "to indicate the reintroduction of a topic" (p. 174).

However, regarding the current study, the data are too limited to make a meaningful interpretation of the subjects' *wa* and *ga* switches. In addition, the grammaticality of example 7 is suspect.

Janakute *and* ja: In her study of Japanese college student CS, Fotos (1994) noted the students' use of *janakute* to signal that the preceding utterance was incorrect and would be repaired. She also noted use of *ja* to attract and focus attention on the utterance to follow. Similar instances were found.

8. T: So you had some time to think about your one minute speech. Okay, now who would like to be the first speaker?
   SS: [laughs]
   Maki: *Ja*, I'll try.
9. Fumi Uh there uh there is there there isn't there isn't uh penici penicillin *janakute* strep strep streptomycin mmm.

Conjunctions: Ss seemed to use conjunctions to help sort out their thoughts and get their messages out. As with tag switches, these may have been used to gain thinking time. Fumi said she wanted to speak English correctly so she was always thinking about whether or not she had made a mistake. Kono said that when something was difficult to express, Japanese seemed to just come out.

10. Fumi: Yes, Switzerland, uh there is there are uh four language in Switzerland, French, eh *eto* German
    Kono: *Doitsu* (Germany). Germany.
    Fumi: Germany
    T: German
    Kono: German
    Fumi: Germany, *sorekara* (and), Germany, German, English and eh, Romansh.[5]
11. Kono: Yes, mm one time is Thursday night, mm *dakedo* (but) mm I have one time *dakedo* (but) mm I *nandakke na* (what was it) mm I hope is two time not not uh another schedule, is one time.

In Example 10, Fumi may be trying to gain time to think or work out the mistake. In Example 11 Kono may be using *dakedo* to get meaning across.

Adverbs and adjectives: Adverb and adjective switches appeared to function as ways to express personal feelings and show solidarity with classmates. In one instance Kono read a sentence and the other students had to guess who it was about.

12. Kono: started working thirty years a...go, go *ga okashii.* (go seems strange)
    T: (laughs)
    Emi: *Ja*, Kono [laughs]. (Well, I guess that's about you.)
    Maki: *Tereteru* [laughs]. (Kono feels self-conscious.)

Maki expressed her feelings about Kono in Japanese. During the interview, Maki said that her first priority was not always to practice English but to communicate with friends.

Fillers: Fumi and Remi made most of the filler switches. During the interview, Fumi said she had not been aware of using *anoo* until asked about it. Remi used the word *nani* (what) as a filler and reported that it had the feeling of "uh" for her.

13. Fumi: K *anoo* Tomohiro Museum uh where where is *anoo*
    Remi: Kusaki Dam? Near the Kusaki Dam.
14. Remi: How was your mother? Walk uh *nani* she took for a walk, uh took walk?

The filler switches may be functioning as a way to gain time and/or to show that the speaker needs help.

Nouns: Nouns were switched frequently. In a conversation about the summer water shortage, Maki explained how she tried to conserve water by giving plants the water she saved from washing rice (*kome*). She ended up using more water because she had to rinse out a rag (*zokin*).

15. Maki: Yes mmm but I I try the uh *kome* rice rice wash a *ko* uh rice. Uh, it's plant [laughs] but a roof roof of the uh wet I [laughs] had uh *zokin*.

For Maki in this instance, the Japanese word, *kome*, may have been more available than the English word, rice. While reporting on her performance in general, Maki said that she knew her ability was not very high and felt impatient while trying to communicate what she really wanted to talk about. She felt she had to explain herself more completely and add more even if it meant using an L1 word.

While talking about her trip to Switzerland, Fumi repeated "glacier lake" in Japanese. She may have thought that it was a difficult term for her classmates and switched to clarify the meaning.

16. Fumi: Switzerland very beautiful country, mountain and lake, eh glacier lake, eh *hyogaka no mizu*, very deep, deep green color.

Fumi said she sometimes uses a Japanese word because she hoped someone would give her an L2 equivalent, i.e. she fished for English words from the class.

Remi, while discussing a similar instance of noun switching (*minami juujisei* for the Southern Cross), said that she switched to Japanese to make sure that her classmates understood what she was talking about.

Nan to iu: Maki made more than half of the *nan to iu* switches. These switches seemed to have been made when the learners did not know the L2 word or could not come up with the item quickly enough.

17. Maki: . . . I tomorrow today morning this morning I take uh I my grandchild take a *boikuen te nan to iu no?* (How do you say nursery school in English?)
18. Emi: Who is play?
Kono: Uh James uh mmm uh James *nandakke naa?* [laughs] Let's see, James what?)

The learners seem to be using these switches to gain time or request help with a word. Variations include *nandakke*, *te iu no*, and *te ieba ii no*.

Phrases and sentences: The Ss used numerous L1 phrase and sentence switches. They sometimes switched to Japanese to speak among themselves, effectively excluding this researcher from the conversation. During one lesson, the women seemed sleepy toward the end so we did some simple bending and stretching exercises to wake up. After the exercises this conversation took place:

19. Remi: *Kyoo wa sugoku tsukareru.* (I'm very tired today.)
SS: Uh, mmm [laughs]
T: Okay, now, okay.
Maki: *Kimari ga ii yoo na ki ga shite, koko de.* (I feel as if this would be a good time to end the lesson.)

Fotos (1994) found instances of CS where feelings are expressed in Japanese and factual information in English in her college student data. Similar instances were collected.

20. Remi: Raise your left hand and stretch.
T: Stretch. That feels good.
Remi: And put it down. Bend your *nani* um bend you *mae.* [laughs] *Zenbu wasuretteru.* (Bend your what um bend forward. I have forgotten how to say everything in English.)
21. Maki: Yes, maybe. [laughs]
SS: [fanning selves, whispering]

Maki: *Nugenai no konna no. Hazukashii.* (I can't take this off. I'm embarrassed.) [laughs] I have a plan.

In example 20 Remi gives commands in English and in 21 Maki answers in the TL, then both switch to the L1 to express feelings of helplessness and embarrassment.

Ni naru *and* ni suru: These intrasentential switches were very simple. They require little grammatical control and they do not violate the equivalence constraint.

22. Maki: Short break *ni suru?* (Shall we take a short break?)
23. Emi: Five weeks *ni naru.* (It will be five weeks.)

Discussion

The use of L1 switches such as tags, conjunctions, fillers, nouns, and *nan to iu* may help learners gain thinking time, smooth the conversation, get important points across, and signal for help. *Ja* and *janakute* switches may function to attract and focus attention on important L2 content and utterances. The subjects may be using interjection, adjective, adverb, noun, and sentence switches to express personal feelings and confirm their solidarity. The learners' CS might be labeled discourse-related switching (Martin-Jones, 1995), "speaker oriented" because it "serves as a resource for accomplishing different communicative acts at specific points within interactional sequences" (p. 99).

The motivations underlying the learners' language switching appear to be similar to some of the factors Gardner-Chloros (1991) gives for French/Alsatian bilingual CS in Strasbourg. Gardner-Chloros (1991) said CS is a kind of compromise

> between the exclusive use of one language and of the other, each with their respective cultural connotations; there are occasions, for example, when it seems too snobbish to speak French but too rustic to speak Alsatian and code-switching provides the solution. (p.184)

> For EFL learners, it may seem too impersonal or difficult to speak English but too unlearner-like to speak Japanese. Thus, they codeswitch.

The motivations Nishimura (1995) outlined for Japanese/English Nisei CS might also be compared to those underlying these subjects' CS. Nishimura (1995) said the Niseis used their mixed variety when speaking to both native Japanese speakers and other Niseis to "reach out" to both types of speakers and "enhance rapport between speaker and the

hearer" (p. 167-169). Japanese EFL learners may want or need to involve both the teacher and each other in communication, or they may be appealing to their dual identities of L1 speaker and L2 learner.

Legenhausen (1991) called CS a mode or register which learners feel

> to be the proper expression of their ambivalent psychological state, pitched between responsibility for their own learning process on the one hand, and their natural inclination to use their L1 in any communicative situation on the other. (p. 71)

The subjects here are in the classroom not just to learn English. They are individuals who appear quite aware of each other and who seem to want to get along. This may explain Fumi's switch:

24. T:	[to Fumi] Comment, question or comment?
Fumi:	Eh, okay, mmm eh, eh, *chotto saki ni, moo jikan ga mottainai.* (Please let someone else go ahead, I'm taking up too much time.)

As Burt (1990) found in her study of learner CS, "it is not always polite to be an extremely conscientious language learner" (p. 34).

It is important for teachers to keep in mind that foreign or second language learners are not becoming monolingual, they are becoming bilingual (Kasper, 1994). Our standards of competency and performance should include bi/multilingual models. There is also much to learn about the teaching and learning processes through the study of CS in the classroom. Martin-Jones (1995) argued that

> a conversational analytic approach to code switching in classroom discourse, grounded in ethnographic observation, can give ... fine grained descriptions of the ways in which teachers and learners get things done bilingually in the classroom. (p. 103)

This study has been limited to learner CS in the classroom. Future studies should look at CS in the interactive discourse of teacher and students.

Notes

1. In the examples, the language other than English is in italics and an idiomatic translation follows in parentheses.
2. See Grosjean (1982) for an introduction to CS from a psycholinguistic perspective.
3. The author's Japanese, Level One on the Japanese-Language Proficiency Test (December, 1992), was considered sufficient for these translations.
4. It should be noted that more information might have been have been gained by looking at switching from both directions.

5. Romansh was recognized as Switzerland's fourth national (though not federally offical) language in 1938.

Ethel Ogane teaches at a private language school and a private university in the Kanto area. She has an M.A. in TESOL from Teachers College, Columbia University and is currently pursuing an Ed.D. in TESOL at Temple University Japan.

References

Appel, R., & Muysken, P. (1987). *Language contact and bilingualism*. London: Edward Arnold.

Auerbach, E.R. (1993). Reexamining English only in the ESL classroom. *TESOL Quarterly, 27,* 9-32.

Blom, J., & Gumperz, J.J. (1972). Social meaning in linguistic structure: Code-switching in Norway. In J.J. Gumperz, & D. Hymes (Eds.), *Directions in sociolinguistics: The ethnography of communication* (pp. 407-434). New York: Holt, Rinehart and Winston.

Burt, S.M. (1990). External and internal conflict: Conversational code-switching and the theory of politeness. *Sociolinguistics, 19,* 21-35.

Cohen, A. D. (1987). Using verbal reports in research on language learning. In C. Faerch, & G. Kasper (Eds.), *Introspection in second language research* (pp. 82-95). Clevedon, Avon: Multingual Matters.

Faerch, C., & Kasper, G. (Eds.). (1983). *Strategies in language communication*. London: Longman.

Fotos, S.S. (1990). Japanese-English code switching in bilingual children. *JALT Journal, 12,* 75-98.

Fotos, S.S. (1994, October). Japanese-English code switching in EFL classrooms. Paper presented at the JALT International Conference, Matsuyama, Japan.

Fotos, S.S. (1995, September). Using Japanese to learn English: Codeswitching in the EFL classroom. Paper presented at the JACET (Japan Association of College English Teachers) Annual Conference, Tokyo.

Gardner-Chloros, P. (1991). *Language selection and switching in Strasbourg*. Oxford: Clarendon Press.

Gardner-Chloros, P. (1995). Codeswitching in community, regional and national repertoires: The myth of the discreteness of linguistic systems. In L. Milroy, & P. Muysken (Eds.), *One speaker, two languages: Cross disciplinary perspectives on code switching* (pp. 68-89). Cambridge: CUP.

Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism*. Cambridge, MA: Harvard University Press.

Gumperz, J.J. (1982). *Discourse strategies*. Cambridge: CUP.

Kasper, G. (1994, May). A bilingual perspective on interlanguage pragmatics. Paper presented at the Tokyo-JALT Annual May Conference, Tokyo.

Kasper, G., & Blum-Kulka, S. (1993). *Interlanguage pragmatics*. New York:

OUP.

Legenhausen, L. (1991). Code-switching in learners' discourse. *IRAL, 29*, 61-73.

Lin, A.M.Y. (1988). Pedagogical and para-pedagogical levels of interaction in the classroom: A social interactional approach to the analysis of the codeswitching behaviour of a bilingual teacher in an English language lesson. *University of Hong Kong, Language Centre, Working Papers in Linguistics and Language Teaching, 11.*

Lin, A. M. Y. (1990). Teaching in two tongues: Language alternation in foreign language classrooms. *City Polytechnic of Hong Kong, Department of English, Research Report 3.*

Lipski, J. M. (1978). Code switching and the problem of bilingual competence. In M. Paradis (Ed.). *Aspects of bilingualism* (pp.250-274). Columbia, SC: Hornbeam Press.

Martin-Jones, M. (1995). Codeswitching in the classroom: Two decades of research. In L. Milroy, & P. Muysken (Eds.), *One speaker, two languages: Cross disciplinary perspectives on code switching* (pp. 90-111). Cambridge: CUP.

McClure, E. (1981). Formal and functional aspects of the codeswitched discourse of bilingual children. In R.P. Duran (Ed.), *Latino language and communicative behavior* (pp. 69-94). Norwood, NJ: Ablex.

Milroy, L., & Muysken, P. (1995). *One speaker, two languages: Cross disciplinary perspectives on codeswitching*. Cambridge: CUP.

Myers-Scotton, C. (1993a). Common and uncommon ground: Social and structural factors in codeswitching. *Language in Society, 22*, 475-503.

Myers-Scotton, C. (1993b). *Dueling languages: Grammatical structure in codeswitching*. Oxford: Clarendon.

Myers-Scotton, C. (1993c). *Social motivations for codeswitching: Evidence from Africa*. Oxford: Clarendon.

Nishimura, M. (1989). The topic comment construction in Japanese-English code-switching. *World Englishes, 8*, 365-377.

Nishimura, M. (1995). A functional analysis of Japanese/English codeswitching. *Journal of Pragmatics, 23*, 157

Poplack, S. (1980). Sometimes I'll start a sentence in English *y termino en Español*: Toward a typology of code-switching. *Linguistics, 18*, 581-618.

Poplack, S. (1988). Contrasting patterns of codeswitching in two communities. In M. Heller (Ed.), *Codeswitching: Anthropological and sociolinguistic perspectives* (pp. 215-244). Berlin: Mouton de Gruyter.

Romaine, S. (1995). *Bilingualism* (2nd ed.). Oxford: Blackwell.

Scotton, C.M., & Ury, W. (1977). Bilingual strategies: The social functions of code-switching. *International Journal of the Sociology of Language, 13*, 5-20.

Tabouret-Keller, A. (1995). Conclusion: Codeswitching research as a theoretical challenge. In L. Milroy, & P. Muysken (Eds.), *One speaker, two languages: Cross disciplinary perspectives on codeswitching* (pp. 344-355).

Cambridge: CUP.
Tarone, E., Cohen, A. D., & Dumas, G. (1983). A closer look at some interlanguage terminology: A framework for communication strategies. In C. Faerch, & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 4-14). London: Longman.

(Received August 15, 1995; revised December 11, 1996)

Perspectives

Japanese and American Television Commercials: A Cultural Study with TEFL Applications

Paul Stapleton
Hokkaido University

This paper suggests that Japanese and American television commercials reflect the values of their respective cultures and are therefore pedagogically useful in the EFL classroom. A preliminary examination of 200 commercials from Japanese and American television programs indicates that there are significant differences in the types of commercials shown, with 54% of the Japanese commercials surveyed appealing to emotion, compared with only 7% of the American, and 66% of the American commercials using logic as the central theme, compared with only 10% of the Japanese. As rich sources of both cultural information and natural language, commercials provide EFL teachers with a means for promoting learner awareness of the values of the target culture through communicative language use.

本論は日本とアメリカのテレビ・コマーシャルがそれぞれの文化を反映しており、したがって英語の授業に有用であると示唆する。日本とアメリカのコマーシャル200点を調査したところ、日本では54％が感情に訴えるものであるのに対し、アメリカでは感情に訴えるものは7％のみであり、中心的テーマとして論理を使うものが66％あった。日本では論理を使ったものは10％のみであった。文化情報と自然な言語の豊かなリソースとしてコマーシャルは学習者にコミュニカティブな言語使用を通して目標文化の価値観に気づかせる手段となる。

Although understanding the values and beliefs of the target culture is of great assistance when learning a second/foreign language, these values and beliefs are often difficult for learners to discover. Therefore any resource which can vividly reveal cultural information should be considered valuable for teaching. This paper suggests that

television commercials are one such resource and examines 200 Japanese and American television commercials to show how they can reflect the values of their respective cultures.

Culture-Based Features of Television Commercials

In the twentieth century, advertising has become a part of the cultural landscape as well as a big business. With prime time commercials in the United States costing in excess of half a million dollars for one minute (Wells, Burnett, & Mociarty, 1995, p. 379), advertisers must know the most effective means to stimulate the viewer to buy the product. This means drawing on input from such diverse fields as psychology, sociology and anthropology–essentially, the culture of the viewer (Wells et al., 1995, p. 163). Advertisers who fail to communicate effectively risk wasting millions of dollars in a matter of minutes and therefore, because such huge sums are at stake, advertisers must ensure that their product, the commercial, produces the desired result.

The Culture-Oriented Model used by advertisers in strategic planning (Wells et al., 1995, p. 752) suggests that although people have common needs, these needs are met in different ways. Thus, it is understood that a commercial which entices an American viewer to buy a product will not necessarily cause a Japanese viewer to react in the same way. An important study conducted in 1990 by Alice, a French advertising agency and IPSOS, a French research group (described in Wells et al., 1995) illustrates this point. Commercials from six European countries were presented 100 consumers from each of the same six countries. The results indicated that country-specific advertising cultures existed. Furthermore, the most competitive commercials in each market were produced by the advertisers who geared their commercial to the specific values of the local culture. Similar to rhetorical patterns (Kaplan, 1966, 1987), it is suggested that commercials reflect culture-specific values and ways of communicating.

The following section presents a classification system for commercials and an examination of television commercials in Japan and the United States. It is suggested that such analysis can reveal important differences in the way Japanese and Americans think and communicate.

The Study

Method

One hundred TV commercials from Japan and the United States respectively were video recorded during February and March, 1996. These commercials were taped from major national television networks from 18:00-23:00. The first 100 commercials recorded on each tape were used for analysis.

Regarding the classification system used to code the recorded commercials, advertising mogul David Ogilvy (1983, p. 103-110) has identified 13 commercial types commonly used in American commercials (see Table 1): humor, slice of life, testimonial, demonstration, problem-solution, talking head (the head and shoulders of a pitch-person selling the product), characters (like the underworked Maytag repairman in US commercials), rational reason why, news, emotion, celebrities, cartoons, and musical vignette. Sommers, Barnes, and Stanton (1995, p. 606) more recently used terms similar to Ogilvy's, indicating that commercials, at least in a broad sense, have not evolved beyond the types that Ogilvy outlined in the early 1980s. Wells et al. (1995, p. 493-495) has classified commercials more broadly, using terms such as storytelling or sight and sound. However, these categories contain the commercial types identified by Ogilvy, suggesting that his classification remains valid today.

Ogilvy's classification (1983) was based on American commercials, and although it was found that the 100 American commercials used in this study fit his typology, it became apparent that the nature of Japanese commercials was somewhat different. Many Japanese commercials were based on appeals to the emotions of the viewers and precise terms specifying the type of emotion appeared wanting. Thus, Ogilvy's list was modified and broken into two broad divisions to emphasize the differences between Japanese and American commercials. The first division, Logic, includes those commercials which provide rational reasons for buying a product. The second division, Emotion, includes commercials which appeal to a viewer's sentiments. As an interesting aside, Ogilvy's, firm, Ogilvy and Mather, failed in Japan both as a joint venture and on its own (Huddleston, 1990, p. 168).

Using the classification system presented below, the commercials were viewed a second time and coded into categories by the author and a second rater. This rater was an American teacher for the American commercials, and a Japanese research assistant for the Japanese commercials. Although the two raters had discussed the evaluation process, classifying commercials according to type was relatively subjective and

Table 1: Commercial Types

Logic (viewer is given a rational argument for buying the product)

1. Explanation – a voice-over describes the merits of the product
2. Slice of Life – two or more characters discuss the merits of the product
3. Testimonial – loyal users of a product testify to its virtues
4. Demonstration – a demonstration of how the product works or why it is good
5. Problem-solution – presents a problem then show how the product solves it
6. Talking head – a pitch-person extols the virtues of the product
7. Price – focuses on the value of the product
8. Comparison – compares the product to the competition

Emotion (viewer is led to associate the product with positive or high quality images)

1. Mood enhancement – creates a positive atmosphere around the product
2. Humor – attempts to amuse the viewer with images, dialogue, sounds, puns
3. Nostalgia – arouses feelings of nostalgia in tandem with the product
4. Celebrity – associates the product with a famous person
5. Musical vignette – a parade of fleeting images played to music
6. Animation – cartoons (almost always used in tandem with other methods)

(adapted from Ogilvy, 1983, pp. 103-110)

disagreement was bound to arise. For the American commercials, agreement between the two raters was 86%, with the first two dozen commercials showing the greatest disparity. Inter-rater agreement was only 75% for the Japanese commercials. This lower figure was probably due to the author's limited Japanese language ability as well as his lack of awareness of Japanese television celebrities.

An additional complication was the fact that many commercials combined elements of several types. For example, a commercial that used humor would also describe some merit of the product being advertised by means of a demonstration or explanation. Consequently, many commercials in the survey were classified as belonging to more than one type, for example, humor/demonstration, and were given two codes. For the Japa-

nese commercials, only 19 were assigned a single code and 81 commercials were placed in two or more categories. For the American commercials, 17 were coded within a single commercial type, while the remaining 83 were assigned two or more codes. Thus, when one-way chi square tests adjusted for continuity (Hatch & Lazaraton, 1991, p. 395) were used as a test of the significance of differences in the types of Japanese and American commercials, the statistical procedure could only be used for the logic/emotion typology, where the items in each cell were independent.

Results

The results of the survey, given in Table 2, agree with previous findings regarding the culturally specific nature of television commercials ("The Enigma," 1993; Huddleston, 1990; Mosdell, 1986; Reed & Kimura, 1990). A significant difference (χ^2crit, 1 *df*, 3.84; χ^2obs. 40.06,

Table 2: Logic-Emotion breakdown of 100 Japanese and 100 American commercials

Heading	Criteria	Japan	US
Logic	Commercials with some element of logic	42%	93%
	Commercials that were completely logic-oriented	10%	66%
Emotion	Commercials with some element of emotion	89%	34%
	Commercials that were completely emotion-oriented	54%	7%

Breakdown of types that showed cultural significance

Type	Criteria	Japan	US
Humor	Commercials with some element of humor	37%	16%
	Commercials that were completely humor-oriented	6%	1%
Celebrity	Commercials with celebrities	42%	3%
Direct Comparison	Commercials mentioning their product's competition	1%	10%
Demonstration	Commercials showing some kind of demonstration	25%	23%
Animation	Commercials with some animation	15%	6%

$p<.05$) was observed between the number of emotionally based commercials for the two cultures, with 89% of the Japanese commercials containing emotional elements, and a full 54% being completely emotional in content. Similarly, significantly more American commercials (χ^2crit, 1 *df*, 3.84; χ^2obs. 50.99, $p<.05$) were logically-oriented; 93% had some elements of logic, and 66% were entirely logical. To explain this difference, three cultural features will be discussed: group vs. individual society, intellectual roots, and the role of silence in communication.

Group versus individual society

Japan is often characterized as a group-oriented society while the United States is said to be based on the individual (Adler & Rodman, 1994, p. 51). It has been suggested that group societies are characterized by homogeneous populations with long, uninterrupted histories, this resulting in what anthropologist Edward Hall calls *high-context* communication (Hall & Hall, 1987, p. 8). Such communication tends to rely on non-verbal contextual elements to deliver a message. These include the way the words are spoken, the gestures and facial expressions that accompany the words (Tenhover, 1994, p. 86), the relationship of the speaker to the listener, and the situation in which the conversation occurs. These non-verbal features of communication are then interpreted by the listener/viewer. In fact, Hall (1983, p. 63) claims that when Japanese are confronted with the deductive methods used in a *low-context* communicative style, where the verbal element is paramount and contextual cues are of much less significance, they may feel as if someone is trying to get inside their heads to do the thinking for them. High-context communication, then, involves letting the communicators use contextual clues to process the meaning of the interaction for themselves (Adler & Rodman, 1994, p. 102).

One example of the high-context nature of Japanese communication can be seen in the thickness of *manga*, Japanese comic books. Because the context of the utterance is such an important part of Japanese communication, *manga* include a multitude of panels which set the mood of the story, taking more pages than what is common in English-language comics (McCloud, 1994, p. 80). Many of these panels are devoid of dialogue and may have no influence on the plot. Rather, they play to the reader's emotions. Comics in English tend to be much thinner because the focus is on the storyline; that is, they are goal-oriented (McCloud, 1994, p. 81).

Extending this idea into the genre of TV commercials, the significantly low number of logic-oriented commercials and high number of emotion-oriented commercials (Table 2) also suggest high-context communication. Japanese commercials tend to take an approach that avoids direct promo-

tion of the product (Huddleston, 1990, p. 181). Instead, they use visual images and oblique references. Overt explanations that logically describe the product, if used, may be secondary. Not surprisingly, this often results in commercials that are enigmatic to foreigners ("*The Enigma*," 1993). However, such failure to comprehend can be expected by non-group members because of their lack of knowledge about Japanese culture.

Consider the following example. During the spring of 1996, a one-minute commercial for Nescafé, a brand of imported instant coffee, was seen on Japanese television. The audio portion of this commercial consisted of music interspersed with a voice-over that lasted for only 13 of the 60 seconds. The voice-over referred to the main figure, a relatively well-known Japanese artist, as a quality artist. Many of the visual sequences showed the main character walking through bamboo forests, or watching cherry blossoms fall. Only fleeting references were made to Nescafé. The viewer was meant to piece the images together to understand that Nescafé was associated with a person of high quality as well as with cultural icons such as bamboo and cherry blossoms which represent the essence of Japan. It has been suggested that Nescafé is one of the most successful foreign products promoted by Japanese advertising (Huddleston, 1990, p. 193), perhaps because of the advertisers' recognition of the Japanese preference for high-context communication.

Interestingly, three of the Japanese commercials surveyed had voice-overs which were almost entirely in English, again indicating the importance of image or mood over logical input. It has been suggested that the English language has become an important cultural symbol; it is not only a system of signs but is a sign in itself (Cheshire & Moser, 1994; Haarman, 1989), evoking the cultures where it is spoken as a native language as well as connotations associated with its use as an international language (Cheshire & Moser, 1994, p. 451). Undoubtedly, only a tiny segment of the viewership of these commercials would be able to understand the meaning of the English. The English content is not the point though; rather, it is the fact that English is being spoken. With the audio portion of these commercials incomprehensible, the average viewer is focused on the visual images. Again this is in tune with high-context communication.

Although Table 2 suggests that Japanese commercials are primarily emotion-oriented and American are logic-oriented, commercials using demonstration seem to stand out as exceptions because they are usually logic-oriented, yet both countries had similar percentages, 25% and 23%. However, demonstration commercials rely mainly on visual images and context, which is characteristic of high-context communication.

In contrast to the high context communication found in Japan, the United States is generally characterized by low-context communication (Adler & Rodman, 1994, p. 50; Hall, 1983, p. 8). Hall suggests that because Americans are members of a young country whose people have diverse backgrounds, they tend to prefer clarity and praise eloquence. With high mobility being a characteristic that dates back centuries to many Americans' European ancestry, the need to communicate with speakers of various languages has been a part of Euro-American culture and it is words alone that determine whether communication is successful (Tenhover, 1994, p. 85). Expressions like "What's the bottom line?" and "Get to the point" reveal the American desire for explicitly worded communication.

American commercials tend to reflect this low-context pattern of communication, with over 90% being logic-oriented and less than 10% being completely emotion-oriented. One American commercial viewed in this study clearly illustrates the value placed upon logical, audio input: The announcer in a car commercial could be heard gasping for air because he was talking so fast, trying to fill every second with information about the advertised car's attributes. A further example of the American desire for information over image is the "infomercial," a special type of long commercial which supplies the viewer with what appears to be a surfeit of information about a product.

Other elements suggested to characterize group culture are the expectation of conformity (Miyamoto, 1994; Mosdell, 1986, p. 19) and shared cultural knowledge and opinion. These elements are manifested in Japanese commercials through the abundant use of celebrities, who appeared in 42% compared with only 3% of the American commercials. For successful context-rich communication, it is considered important to keep the potential customer informed about the latest trends. Celebrities often act as opinion leaders who are on the cutting edge of what is in style. In this sense, celebrities are guides for the public to follow. The commercials for a number of products featuring the Los Angeles Dodgers pitcher from Japan, Hideo Nomo, in 1995 and 1996 are an example of the celebrity trend-setter genre.

On the other hand, American commercials suggest cultural values which are quite different. America describes itself as "the land of the free" and commercials such as one for Nike shoes urging viewers to "Just do it." A Ford commercial states, "We stand out." Such commercials reflect an American preference for individualism and the desire to do what one chooses without regard for what others may think. The low frequency (only 3%) of celebrities in American commercials is in keeping with Ogilvy's claim (1983, p. 109) that viewers tend to remember the

celebrity but forget the product being endorsed. Furthermore, other researchers (Wells et al., 1995, p. 691) have questioned whether American viewers really believe that the celebrity uses the product.

Perhaps in an individual-oriented society people do not wish to be followers, so seeing a celebrity endorse a product is not necessarily a guarantee that the product will sell. Naturally, there are exceptions to this general rule. In recent years, market research results have shown that sport and pop icons like Michael Jordan, Michael Jackson, and Madonna have been able to transcend the celebrity taboo.

Intellectual Roots

Although both Japanese and Euro-American cultures have many and varied intellectual influences, noted scholar and former ambassador to Japan Edwin O. Reischauer (1988, p. 204) suggested that the most profound influence in Japan has been Confucianism. In contrast, American intellectual traditions have been shaped by thought from ancient Greece. Of course, such a characterization is greatly oversimplified but some values of Confucianism such as harmony and modesty are apparent in Japanese commercials, whereas manifestations of Greek philosophy in American commercials are evident in their logically-based presentation of information and their rhetorically-dominated nature.

As mentioned, almost 90% of the Japanese commercials had some element of emotion. Emotional appeals have no rational influence on the purchase of a product; instead, they attempt to touch the viewer with feelings such as humor and nostalgia which incite positive feelings towards the product. Emotionally-oriented commercials avoid boasting about a product's merits or directly comparing the product with its competition. Advertisers evidently believe that in this way competitive harmony is maintained and that the viewer is not driven to feel uncomfortable about immodest claims. In the Confucian way of thinking, claiming that one's product is better than a competitor's is not only immodest, but it is meaningless as well because it is up to the consumer to decide which product is best (Mosdell, 1986, p. 10). Interestingly, only one direct comparison commercial was found in the Japanese sample; it featured a Pepsi product opposite a Coca-Cola product. This almost complete lack of openly competitive advertising in Japan, and the frowning upon it by some authorities (Huddleston, 1990, p. 161), suggests that the Confucian ideals of harmony and modesty are exemplified in many Japanese commercials.

Although it has been noted that emotion-oriented advertising is both common and effective in the United States (Ogilvy, 1983, p. 109; Wells et

al., 1995, p. 280), most of the American commercials in this study tended to be based on elements of logic and rhetoric. The need for information to support one's idea or decision is a cultural tradition derived from Greek philosophy. Plato advocated the dialectic, or debate, in which conflicting information is used in an effort to reach a conclusion. Advertisers recognize that many American viewers require such information in order to make a decision (Adler and Rodman, 1994, p. 102). This suggestion is somewhat supported by the data presented here. Ten percent of the American commercials used direct comparison, where the advertised product was presented beside its competitor and claimed to be superior, compared with only 1% of the Japanese commercials. One American commercial advertising a pain medicine actually lined up its product beside the three other market leaders, which could be identified by pill shape and logo. The pitch-person, while identifying one of the competitors, said that his product "just works better." Such immodest claims are commonly seen in American commercials. In the absence of a deep-down belief in Confucian harmony, Americans appear to respond positively to competition (Stewart and Bennett, 1991, p. 79).

Silence: In *The Silent Language* (1959), Edward Hall presented the idea that perceptions of time and space differ from culture to culture. It has been suggested (Nanda, 1994, p. 81) that Americans consider time and space as things to be filled. Time must be filled by some activity, and if space is not filled, it is thought of as empty or wasted (Nanda, 1994, p. 81). For Japanese, however, both space and time have intrinsic meaning. If there is silence during a conversation, the silence does not indicate emptiness; rather, it communicates a message.

In their use of sound, commercials can convey a sense of how a viewer understands aural messages. A common feature of Japanese commercials is their relative lack of a voice-over (Huddleston, 1990, p. 161). Instead, there is often simply music or sometimes silence. The value of silence in Japan can be related to Zen Buddhism, which can be loosely described as a way of learning which requires the student to master a skill by following formalized rituals (Hori, 1996, p. 26; Gotz, 1988). Herrigel (1953) writes that a Zen master shuns "long-winded instructions and explanations, (and the Zen teacher) contents himself with perfunctory commands and does not reckon on any questions from his pupil" (Herrigel, 1953, p. 45). Within the Zen tradition, words play only a minor role while silent communication is the essence of learning. In contrast to American culture, where stress is placed on using words to convey meaning (Stewart and Bennett, 1991, p. 157), Zen requires the

student to eliminate rational thought and operate on a different plane. It should be noted, however, that although silence is an integral part of Zen, there are other explanations for the value placed on silence in Japanese culture (Matsumoto, 1988). Silence played an important role in feudal Japan, where subordinates were mostly silent in the presence of their superiors in recognition of the strict social hierarchy. Likewise, Japanese often speak of *isshin denshin* (Matsumoto, 1988, p. 48) or "heart to heart communication," where words are unnecessary because both speaker and listener are members of a tightly-knit, high-context culture and can understand non-verbal messages.

Regardless of the exact nature of its source, silence in Japan may carry deep feelings (Hall, 1983, p. 99) or may simply give listeners a chance to gather their thoughts or reflect on what has just been said. In Japanese commercials, silence is used to build a positive atmosphere around a product and Japanese viewers intuitively understand such use. In contrast, Levine and Adelman (1993, p. 71) note the American tendency to fill every space in a conversation with words. For Americans, silence is associated with awkwardness, misunderstanding, tension, shyness, disagreement (Tenhover, 1994 p. 156) or "dead air" (Hall, 1983 p. 99), all generally negative qualities. In English, silence is contrasted with eloquence, a highly desired quality since the time of the ancient Greeks. In the fourth century B.C. Isocrates described speech as follows:

> By speech we refute the wicked and praise the good. By speech we educate the ignorant and inform the wise. We regard the ability to speak properly as the best sign of intelligence, and truthful, legal and just speech is the reflection of a good and trustworthy soul . . . . If I may sum up on this subject, we shall find that nothing done with intelligence is done without speech. (Adler & Rodman, 1994, back cover)

The predominance of logic-oriented commercials with a strong rhetorical base serves as testimony to Isocrates' quote. The Japanese Nescafé commercial, whose minimal voice-over hardly mentioned the product, and the American Ford commercial, where the announcer was talking so fast his inhaling could be heard, are representative cases of the contrasting use of sound for communication in Japanese and American commercials.

Commercials as a Teaching Resource

While cultural icons such as food, sports, and even TV commercials are considered the exposed part of culture, underlying themes such as group versus individual values, intellectual roots, and silence as a form of com-

munication are suggested to be part of culture's hidden nature (Levine and Adelman, 1993, p. xviii). This covert culture is very important for foreign language learners. However, unlike the surface or overt culture, which includes the language functions familiar to all language teachers, covert culture is more difficult to teach. It is like the proverbial iceberg, with only a small portion available for easy analysis. Teachers can explain the deeper values that govern the behavior of those who speak the target language. However, such explanations often remain inadequate because many students may be only vaguely aware of the deeper cultural values that govern their *own* behavior. Therefore, any pedagogical resource that can reveal to learners in a vivid way their own covert culture as well as the culture of the target language should be utilized.

TV commercials from the target culture serve as one such resource because of their pervasiveness, their meaning-focused use of the target language, and their rich cultural content. The data from the present study suggest that many commercials reflect important communicative values. However, most students are unlikely to be aware of the covert culture presented by commercials from their own culture, nor are they likely to be able to understand how commercials from America reflect its covert culture. Therefore, contrasting Japanese and target culture commercials as a communicative classroom activity can be illuminating to students who have only seen the surface message that a commercial displays. When introducing commercials to students at lower proficiency levels of the target language, teachers can begin by explaining some of the major cultural themes discussed above. Then they can show commercials from the learners' culture and the target culture which reflect these themes as transparently as possible. As a second step, or at higher proficiency levels, teachers can show commercials from both cultures and ask students to find deeper cultural values on their own, in an effort to develop critical thinking skills. (See the Appendix for a more detailed lesson plan.) To see familiar images in a completely new light can often leave a powerful impression on students.

Paul Stapleton teaches English at Hokkaido University.

References

Adler, R.B., & Rodman, G. (1994). *Understanding human communication*. Fort Worth, TX: Harcourt Brace.

Cheshire, J., & Moser, L-M. (1994). English as a cultural symbol: The case of advertisements in French-speaking Switzerland. *Journal of Multilingual and Multicultural Development, 15*, 451-470.

The enigma of Japanese advertising. (1993, August 14). *The Economist*, p. 61-62.
Gotz, I. L. (1988). *Zen and the art of teaching*. New York: Wilkerson Publishing Co.
Haarman, H. (1989). *Symbolic value of foreign language use*. Berlin: Moutin de Gruyter.
Hall, E.T. (1959). *The silent language*. Garden City, NY: Doubleday & Co.
Hall, E.T. (1983). *The dance of life*. New York: Anchor Books.
Hall, E.T., & Hall, M.R. (1987). *Hidden differences*. New York: Anchor Books.
Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. New York: Newbury House.
Herrigel, E. (1953). *Zen in the art of archery*. New York: Random House.
Hori, G.V.S. (1996). Teaching and learning in the Rinzai Zen monastery. In T. Rohlen and G. LeTendre (Eds.), *Teaching and learning in Japan* (pp. 20-49). Cambridge: Cambridge University Press.
Huddleston, J. N. (1990). *Gaijin kaisha (Foreign companies)*. Tokyo: Tuttle.
Kaplan, R. (1966). Cultural thought patterns in inter-cultural education. *Language Learning, 16*, 1-20.
Kaplan, R. (1987). Cultural thought patterns revisited. In U. Conner and R. Kaplan (Eds.), *Writing across languages: Analysis of L2 text* (pp. 9-12). Reading, MA: Addison-Wesley.
Levine, D.R., & Adelman, M.B. (1993). *Beyond language*. Englewood Cliffs, NJ: Prentice-Hall Regents.
Matsumoto, M. (1988). *The unspoken way*. Tokyo: Kodansha.
McCloud, S. (1994). *Understanding comics: The invisible art*. New York: Harper Collins.
Miyamoto, M. (1994). *Straitjacket society*. Tokyo: Kodansha.
Mosdell, C. (1986). *The mirror makers*. Tokyo: Macmillan.
Nanda, S. (1994). *Cultural anthropology*. Belmont, CA: Wadsworth.
Ogilvy, D. (1983). *Ogilvy on advertising*. New York: Multimedia.
Reed, G. & Kimura, S. (1990). *CM watching*. Tokyo: Macmillan.
Reischauer, E.O. (1988). *The Japanese today*. Tokyo: Tuttle
Sommers, M., Barnes, J., & Stanton, W. (1995). *Fundamentals of marketing*. New York: McGraw-Hill Ryerson.
Stewart, E.C., & Bennett, M.J. (1991). *American cultural patterns*. Yarmouth, ME: Intercultural Press.
Tenhover, G.R. (1994). *Unlocking the Japanese business mind*. Washington, DC: Transemantics.
Wells, W., Burnett, J., & Mociarty, S. (1995). *Advertising: Principles and practice*. Englewood Cliffs, NJ: Prentice-Hall.

(Received November 3, 1996; revised December 11, 1996)

Appendix: Sample Lesson Plan for Teaching Cultural Values using TV Commercials

The following is a description of a 90-minute introductory EFL class in Japan using commercials to teach cultural values. Because many of the ideas presented are likely to be new to the students, much of the lesson is in the form of a lecture or demonstration.

Part One: As an introductory activity, the teacher asks students to write a definition of the word "culture." If this is too difficult, they can be asked to make a list of things which represent Japanese culture. After eliciting answers, the teacher draws an iceberg on the board and writes the students' answers in the appropriate sector of the iceberg, overt for the visible tip of the iceberg and covert for the submerged portion. To represent overt culture, students often refer to art, music, food, fashion, etc. The teacher should explain that covert culture, including communication style, beliefs, attitudes, values, and perceptions (Levine & Adelman, 1993, p. xviii), is the foundation upon which much of overt culture rests.

Part Two: As this is an introductory lesson, only one covert theme should be discussed. It is suggested that "group versus individual culture" is the most apparent when comparing Japanese and American commercials and, therefore, the most easily taught. Some simple examples of how group and individual cultures are manifested in daily life in Japan and the United States should be provided. For example, it could be pointed out that Japanese infants usually sleep in the same room as their parents for several years before getting their own room (Reischauer, 1988, p. 144), and so develop a stronger feeling of being a group member. American children, on the other hand, usually have their own room shortly after birth (Levine and Adelman, 1993, p. 172), a practice which instills independence from an early age.

The students then receive Worksheet 1, which presents the theme, outlines the thematic characteristics of the cultures, and lists the way in which these characteristics might be manifested in commercials. As the concepts are difficult, it is suggested that the worksheet be bilingual. The teacher reviews the contents of the worksheet with the students and discusses the concepts. The students will use this worksheet to help them understand commercials.

Part Three: The teacher shows two or three representative commercials, pointing out the cultural characteristics revealed. The students use the written information from Worksheet 2 to guide their analysis.

Part Four: The teacher shows several more commercials from each culture and the students work individually or in groups to analyze the commercials' representative characteristics, then complete Worksheet 2.

Worksheet 1

Theme	Characteristics	CM Manifestation
Group society (Japan)	high-context communication modesty harmony village culture strict code of conformity inspires need to escape desire to conform	fewer verbal explanations few hard sell cms few direct comparison cms no need to explain animation many celebrities (show viewers how to be part of the group)
Individual Society (the United States)	low- context communication focus on clarity harmony is less important competitiveness is good	explanation-oriented (using words) information oriented boastful CMs / direct comparison CMs direct comparison CMs focus on price

Note: CM, from "commercial message," means "commercial" in Japanese.

Worksheet 2

Commercial	CM's Characteristics	Cultural Significance
car (US) Part 3	focus on price full of explanations mentions competitors	competitiveness is good low-context communication harmony is less important
coffee (Japan) Part 3	mostly silent few references to product celebrity	high-context communication modesty / vagueness desire to conform
Part 4:	*(Completed by students)*	*(Completed by students)*
Part 4:		
Part 4:		

Reviews

Multilingual Japan. John C. Maher and Kyoko Yashiro (Eds.). Clevedon, UK: Multilingual Matters Ltd., 1995. 164 pp.

Reviewed by
Steve McCarthy
Kagawa Junior College

This collection of papers demonstrating the actual linguistic diversity of Japan is a special double issue of the *Journal of Multilingual and Multicultural Development* (1995) in book form. It complements an earlier collection of papers describing the diversity of Japanese society (Maher & Yashiro, 1991). Now they have compiled a work more comprehensive in scope and more representative in content.

Multilingual Japan is a deceptively thin volume, dense with historical information, linguistic data, critical observations, and references for further exploration. Following the introduction come papers on the Ryukyuan, Ainu, Korean, and Chinese languages in Japan, loan words from English, returnees, and bilingualism in international families. Moreover the sole official language, Japanese, is involved with all the rest.

In the introduction, Maher and Yashiro survey the linguistic diversity of this archipelago. Yet to admit this officially would open the door to the more politically dangerous recognition of cultural diversity. They debunk the notion of "the Japanese," arguing that the inhabitants of these islands must be seen anew as just people, without the overlay of myths and stereotypes from second-hand accounts.

The authors discuss the development and standardization of the Japanese language, versus its dialects, sign language, and various minority languages. They deconstruct ideologies of monolingual-monocultural homogeneity and harmony as invented traditions. They see Nakasone's denial of the existence of minorities Japan as a symptom of the statism that, since the Meiji Era, has suppressed both minority aspirations and the sociolinguistic frameworks through which language diversity could be investigated.

The first paper on a specific language, "Ryukyuan: Past, Present and Future" by Akiko Matsumori, makes extensive references to vernacular

research. "Ryukyuan" is preferred to describe the language group spoken in the formerly independent kingdom of Ryukyu, today's Okinawa Prefecture and some islands in Kagoshima Prefecture.

Matsumori details the history and geography of the Ryukyus in addition to analyzing the relationship between Ryukyuan and Japanese. Ryukyuan dialects are related to Japanese and have provided some diachronic clues in reconstructing the elusive history of Japanese. Mastumori observes that Ryukyuan is commonly called the Okinawan dialect of Japanese for reasons more political than linguistic.

In a typical case of language shift, Okinawans below the age of 40 are losing their Ryukyuan fluency and almost everyone speaks standard Japanese or rather interdialects resulting from interference during accommodation. Okinawans themselves have embraced language standardization to the detriment of local dialects, while schools have been draconian in stigmatizing non-standard Japanese usage. Ryukyuans have often been forced to change their social identity, to emphasize their common heritage with mainland Japan either in preference to American rule or for economic reasons.

"English in Japanese Society: Language within Language," by Nobuyuki Honna, does not deliver the sociolinguistic analysis promised, but does provide a valuable introduction to loan words from English. The strength of the paper is a taxonomy of seven types of borrowing patterns that involve semantic and/or structural changes. Since daily Japanese speech includes about 13 percent loan words, mostly from English, knowledge of loan words is necessary if EFL teachers in Japan are to develop strategies toward the variety of English with which the students were raised.

Honna may be listening too much to purists, though, when he writes that loan words alarm many people. An Asahi Shimbun ("*Honsha*," 1996) poll found that among the 77 percent who feel that Japanese usage is degenerating, only 6 percent cited an excess of foreign words, while 28 percent blamed youth slang. Yet Honna makes the redeeming observation that, lest people see compulsory English education as a failure, it has borne fruit by enriching the Japanese language.

"Bilingualism in International Families," by JALT Bilingualism N-SIG co-founder Masayo Yamamoto, summarizes research on bilingualism in Japan along with her own survey findings. She confines the data in this paper to families with one English and one Japanese native-speaking parent. She explains the choice of English between spouses is due to the greater English proficiency of the native Japanese speakers in most cases.

Yamamoto explains that in language use from parents to children at home, the L1 is used most often with one's children for emotional bond-

ing or to consciously impart the language. The force of the societal language, however, is such that more Japanese is heard from the children than is spoken by the parents to them. Bi-literacy is difficult to attain, with bedtime reading by the minority language parent a necessary but not sufficient condition. Physical or linguistic conspicuousness in Japan can result in children resisting English to minimize their differences from the norm. But Yamamoto concludes, those who do become bilingual are generally admired.

Turning from the conspicuous to the partly submerged minorities, Maher and Kawanishi co-author "On Being There: Koreans in Japan." They recount the colonial history of forced labor that resulted in a million residents of Korean heritage. As Korean-medium schools were forbidden among those who stayed after the war, Korean language proficiency among the second, third, and fourth generations is rapidly declining. Today the *Soren* (North Korea-affiliated) and *Mindan* (ROK-affiliated) organizations operate school systems with a bilingual curriculum in Korean and Japanese.

As Korean schools are not accredited, 86% of Korean students attend Japanese schools in order to have any chance of entering national universities. However, many Korean-Japanese attending Japanese schools also receive materials on ethnic education which lovingly portray the culture of the homeland. One text by *Mindan* exhorts all Koreans in Japan to have ethnic consciousness, to live in dignity, and to be true internationalists *(kuk'chae'in = kokusaijin)*. Soren textbooks, on the other hand, tend to be more ideological, singing praises of the North Korean leaders. The lack of consensus among ethnic Koreans also appears in the contentious issue of maintaining Korean names or not in face of the mainstream society.

In "The Current State of the Ainu Language," Joseph DeChicchis's 164 references in several languages illustrates the extensive research on this minority and its language, both termed "Ainu." The number of officially registered Ainu is only 24,000, a result of their historical experience as a downtrodden and partially assimilated minority.

Ainu representatives' petition to the United Nations to recognize them as an indigenous minority treated unjustly, led to a bizarre government statement that the Ainu were Japanese. Similarly, Japanese scholars tend to emphasize Ainu-Japanese linguistic similarities. On the other hand, early reports creating the image of Ainu as Caucasian exaggerated their differences from Japanese. DeChicchis speculates that the Ainu language is non-Altaic but with much Altaic vocabulary, plus later borrowings from Japanese and the northern Asian languages of peoples they contacted.

John Maher then sheds light on "The *Kakyo:* Chinese in Japan." There are at least 50,000 stable residents of mostly inner-city communities using Cantonese and some Mandarin along with JSL. In the second to fourth generations there is a trend toward dominance in Japanese. Post-War *Shin-Kakyo* tend to start with Taiwanese or other dialects.

Another nearly 50,000 speakers of Chinese are neither well-established nor are they called *Kakyo.* Since 1980 there has been an influx of Mandarin-speaking students and laborers speaking various dialects, JSL, and pidgin. In addition, there are returnees from China wishing to be repatriated with their Japanese families.

Schools for *Hua-chi'iao (Kakyo)* exist in the port cities of Tokyo, Yokohama, Kobe and Nagasaki. In Yokohama with 300 pupils from kindergarten through middle school, instruction is in Chinese and Japanese by mostly Japan-born Chinese. Some Japanese children attend as well, analogous to those sent to English- or French- language international schools. At the Tokyo Chinese School, from primary school on, English further augments Mandarin Chinese and Japanese instruction. Maher thus regards the Chinese schools as a model for bilingual education in Japan.

In "Japan's Returnees," Kyoko Yashiro notes the shift from regarding returnees as problem children in need of re-acculturation to a valuable human resource in the search for internationalization. University-age returnees receive privileged quotas at many prestigious universities and enjoy an advantage in being hired by big businesses that send employees abroad. Various government and private sector organizations support them or their networking among each other. Very few returnees now have serious problems of linguistic or cultural readjustment.

However, while Japanese schools have been set up abroad to maintain L1, L2 maintenance has been neglected in Japan, particularly by public schools. Yashiro refutes each rationalization for this neglect. Her surveys of *kikokushijo* have shown that over 90% wish to maintain their L2, virtually all who have anything significant to maintain, regardless of the second language. Returnees thus warrant L2 maintenance as agents of internationalization and diversity in Japanese society.

A weakness of this collection is the lack of final editing, for which the publisher must bear some responsibility. The introduction is strident in tone and a bit disorganized in its laudable attempt to cover disparate issues not treated in the other papers. The typos may unfairly discourage readers from continuing to papers by non-native writers of English though these present relatively few obstacles to understanding.

The collection represent a reliable sociolingusitic study for scholars abroad, while language teachers in Japan can derive applications from a

deeper understanding of our social context. Teachers can combat the unjust portrayal of Japanese students as a monolith and avoid blunders with submerged minorities by learning the variegated particulars beneath the ideology of sameness. For L2 pedagogy, the book suggests ways loan words could be a valuable aid to study, while for the hundreds of JALT members with international families, it provides the sketch of a road map for bilingual child-raising. The more deeply teachers are committed to Japanese society, the more useful this book will be.

References

"*Honsha teiki yoron chosa: Kokumin no ishiki wa*" ("Asahi Shimbun public opinion poll: Citizens' awareness"). (1996, January 1). *Asahi Shimbun*, p. 20.

Maher, J. & Yashiro, K. (Eds.) (1991). *Nihon no bairinguarizumu (Japan's Bilingualism)*. Tokyo: Kenkyusha.

Oxford Advanced Learner's Dictionary, Fifth Edition, Oxford: Oxford University Press, 1995; x + 1,428 pp.

Reviewed by
Brian C. Perry
Nagoya University

The fifth edition of the *Oxford Advanced Learner's Dictionary* (OALD) is the latest work originating from A.S. Hornby's pioneering *Advanced Learners' Dictionary*, first published in Japan in 1948. Since then monolingual dictionaries, written specifically for non-native learners of English (learners' dictionaries), have become firmly established as a reference for intermediate and advanced learners. For such learners, there is general agreement that the benefits of using a dictionary in the target language more than compensate for any difficulties in understanding the definitions (Hartmann, 1992, p. 153). The fifth edition claims to have 65,000 definitions (with 2,300 being new words and meanings), 90,000 examples and 1,700 illustrated words. Included are 16 language study pages and 10 appendices. The smaller compact version, designed more for Japanese learners, differs from the standard version in that it has a soft cover and comes with headwords split into syllabic divisions.

What is different about the fifth edition is that for the first time it is based on corpora of present-day written and spoken English: the 100-million-word British National Corpus and the 40-million-word Oxford American English Corpus. The lexicographical trend towards corpus-based learners' dictionaries, which began with the *Collins cobuild English Dictionary* (Sinclair, 1987), is of immense significance for two reasons. First, compilers can now scientifically analyze their data banks of contemporary language to select the most frequently used words for inclusion in their lists of, (a) headwords, and (b) defining vocabulary. Second, the example sentences–a key feature of any learners' dictionary—can now be based on those recorded as having been actually used. They can also be chosen to contain the word in question in its most frequently found collocations–and therefore be more pertinent for the user.

The layout of the OALD is clear and concise. Headwords are in large bold type with examples in italics. Idioms and phrasal verbs are also marked in bold at the end of each entry. Standard IPA phonemic transcriptions are given along with a pronunciation guide. American spellings, meanings and pronunciations are shown where required. Most learners' dictionaries emphasize their total number of entries, yet learners, unlike native speakers, tend to be more interested in the basic stock of words. Here the OALD seems to have got the balance about right. Its 65,000 entries are enough to cover the needs of almost all advanced learners whilst at the same time being reasonably compact. In comparison, the 100,000-word *Cambridge International Dictionary of English* (CIDE) (Proctor, 1995) is too heavy and bulky to be easily portable.

Hartmann (1992, p. 153) and others have summarized what the main design features a learners' dictionary should have. I will examine the OALD in the context of the three most important (adapted from Hartmann's summary), namely: 1. the definitions are geared to the more limited vocabulary of the foreign learner; 2. collocational detail is provided, usually by example sentences; and, 3. grammatical coding is detailed and specific and stylistic information is given, typically by usage labels.

Definitions: Entries in the OALD are defined using a 3,500-word core vocabulary that includes both headwords and derivatives. In contrast, a rival dictionary, the CIDE, claims a defining vocabulary of only 2,000 words. But these are only headwords, and their derivatives must be added. So, for example, one finds that *absorb* is included in both lists, but the CIDE by default also includes *absorbent, absorbency,* and *absorption*—none of which are likely to be considered common words for a learner. Furthermore, the definitions in the OALD rarely stray from the defining vocabulary so they may indeed be more restricted than other learners' dictionaries. Whether

this is beneficial for the learner has been challenged by some, notably Carter (1987), who believe that limiting the vocabulary used in definitions is unnatural. Clearly there is a trade-off between accuracy and understandability. The OALD tends towards the latter.

Collocational Detail: The OALD does not, unlike the COBUILD, stick rigidly to the sentences thrown up by the corpus, but rather uses them as a basis for its examples. In this respect, the OALD comes down on the side of pedagogy in the long-running debate between those who have argued for authentic examples (Sinclair, 1987, p. xv), and those who have argued for pedagogically-based examples (Hausemann & Gorbahn, 1989, pp. 45-47). If one looks up the word *drainage* in the COBUILD we get the authentic but rather strange-sounding *line the pots with pebbles to ensure good drainage.* The corresponding example in the OALD is the short and concise *land drainage schemes.* Short phrases such as this are often found in the examples, particularly for less common words. On one hand they highlight a frequently found collocation and save space; on the other, they are too short to truly represent authentic English.

Grammatical & Stylistic Information: One of the key uses of a learners' dictionary, such as the OALD, is to help learners to encode the language for writing. It follows then that lexical information on its own would seem to be insufficient. A learner is also interested in how to put lexical units together in a way that is not only grammatically accurate, but is also stylistically appropriate. For individual entries the OALD adopts a system of coding which is clear and informative. There are 220 notes on usage under individual entries. Furthermore, a new system of simplified verb codes, (e.g. [Vnn] = verb + noun + noun) appears before the examples that illustrate them. Usage labels are precise. The entry *byte* for example, is labeled "computing" which is more detailed than "technical" or "specialized" found in the *Longman Dictionary of Contemporary English* (LDOCE) (Summers, 1995) and CIDE respectively.

It would also make sense for learners' dictionaries to include a coherent description of the grammar of English in a systematically ordered component (as called for by Lemmens & Wekker, 1991). However, the displaying of general grammar is patchy. Apart from the appendix on irregular verbs, syntactic information is contained within the misleadingly named study pages. As Walter (1996, p. 358) observes, there is little in the way of cross-referencing between them and the main body of the dictionary. If one looks up the word *idiom,* there is nothing to link it to a well-written two-page section supported by examples headed *idioms* in the study pages.

In addition, the OALD contains illustrations that are sensible and not overused, along with a number of color maps and pages presenting

information on such things as the British and American Constitutions. Illustrations appear to follow the rule, "all items illustrated should be large enough to see and labeled in such a way that there is no likelihood of confusion between items" (Nesi, 1989, p. 133).

At the back there is a comprehensive set of ten appendices which take only 35 out of the 1,428 pages. They are clearly an efficient way of presenting lexical information. Take, for example, appendix 6, "Ranks in the armed forces." It would be long-winded and clumsy to fully define *captain* in the standard way. An accurate definition would have to not only state that a *captain* is a commissioned officer, but also explain its rank for each military service, and for both British and American armed forces. In tabular form one can immediately see its place in each hierarchy of ranks.

The most serious shortcoming of the OALD is the lack of information given about the frequency of usage for a particular word. Such information is especially helpful for non-native learners because it can enable them to give a word an appropriate priority for learning and production. This is particularly true for learners who are studying for exams which specify the number of English words to be mastered. It is somewhat disappointing then to find only a list of its 3,500-word defining "core vocabulary" as an appendix. The core vocabulary is of course not necessarily the same as the 3,500 most frequent words. Words such as *adjective* and *grammar* are only included for the reason that they are useful in dictionary definitions. Moreover, whether a word is in this list is not indicated in the in the main body of the dictionary. To find out, the word has to be looked up again in the list at the back. In comparison both the LDOCE and the COBUILD display the degree of frequency annotated next to the headword, with the LDOCE also showing it for both speech and writing. Looking up the word *insurance* in the OALD, one finds that it is included in the appendix of core vocabulary—the only information about frequency given. However, in the LDOCE it is marked as "S3 W2," meaning that *insurance* is ranked within the first 2-3000 words for speech but within 1-2000 in writing. The OALD is one step behind its rivals in this area.

It would be reasonable to assume that most learners require a dictionary which for them is concise, easy-to-use and above all portable. For them the fifth edition of the OALD is a first-rate learners' dictionary. Despite its shortcomings it broadly fulfills the main functions of a learners' dictionary which is to help non-native learners encode and decode language in a relevant, efficient and accurate manner. For some learners, however, the OALD will not be suitable. Those who require a weightier

desktop reference would probably fare better with the CIDE. Those who attach importance to knowing the relative frequency of entries should consider the LDOCE or COBUILD. And finally, those who require a full description of the grammar of English in the same book will have to wait until a dictionary publisher has the foresight to include one.

References

Carter, R.A. (1987). *Vocabulary: Applied linguistics perspectives*. London: Allen and Unwin.

Hartmann, R.R.K. (1992). Lexicography, with particular reference to English learners' dictionaries. *Language Teaching, 25*, 151-159.

Hausemann, J.H., & Gorbahn, A. (1989). COBUILD and LDOCE II: A comparative review. *International Journal of Lexicography, 2*(1), 45-56.

Lemmens, M., & Wekker, H. (1991). On the relationship between lexis and grammar in English learners' dictionaries. *International Journal of Lexicography, 4*(1), 1-14.

Nesi, H. (1989). How many words is a picture worth? A review of illustrations in dictionaries. In M.L. Tickoo (Ed.), *Learners' dictionaries: State of the art* (Anthology Series, 23) (pp. 124-134). Singapore: SEAMEO Regional Language Centre.

Proctor, P. (Ed.). (1995). *Cambridge international dictionary of English*. Cambridge: Cambridge University Press.

Sinclair, J. (1987). *Collins cobuild English language dictionary*. London: Collins.

Summers, D. (Ed.). (1995). *Longman dictionary of contemporary English* (3rd ed.). Harlow: Longman.

Walter, C. (1996). Review of ELT dictionaries. *ELT Journal, 50*(4), 356-359.

Understanding Communication in Second Language Classrooms. Karen E. Johnson. Cambridge: Cambridge University Press, 1995. xv + 187 pp.

Reviewed by
Damian Lucantonio
The Japan Times

Understanding Communication in Second Language Classrooms is a book for teachers and teacher educators interested in the role of classroom-based research in second language education. The author, Karen E. Johnson, describes the nature of communication in second language classrooms so that more effective teaching strategies can be implemented and draws on data from actual classroom discourse to support the ideas that emerge. This allows her to provide a detailed account of the nature of communication in second language classrooms and to examine the implications of this for teaching and learning.

Her purpose is to describe how the patterns of communication are established and maintained in second language classrooms, and how these effect classroom performance, classroom learning, and language acquisition (p. 3). Johnson demonstrates that classroom communication does in fact have a distinct, if complex and variable, system. For example, the theoretical framework that underpins the book (p. 8) acknowledges the roles of both teachers and learners in the moment-to-moment activities that occur within classrooms. The meanings communicated are shaped by the actions and perceptions of teachers and students, as well as by the social, cultural, and institutional contexts within which classrooms exist. These actions and perceptions can differ widely from teacher to student, student to teacher, and student to student, which in turn can affect student participation, performance, and achievement.

The framework for understanding communication in second language classrooms used in the book is adapted from a model of communication and learning from British researcher Douglas Barnes (1976). He posits that students are active receivers of knowledge and believes that classroom learning is a negotiation between teachers' meanings and students' understandings. Furthermore, he argues that classroom communication occurs along two major dimensions. The first is the moment-to-moment actions and interactions that occur in classrooms. The second represents what teachers and students bring to the classroom in terms of knowledge and attitudes. The two dimensions interrelate to shape the classroom communication that takes place (p. 7).

The book has three main sections. Part 1, chapters 1 - 5, discusses each component of the theoretical framework and evaluates the contribution of each to classroom communication. Chapter 1 surveys the framework and illustrates the communicative demands placed on teachers and students in second language classrooms. Chapter 2 focuses on teachers' control of the patterns of classroom communication. For example, Johnson analyzes the IRE teaching sequence (the initiation act, the response act, and the evaluation act) (p. 17) which underlies the structure of many classroom lessons (Sinclair & Coulthard, 1975). The chapter examines the ways teachers use language to control the structure and content of classroom activities and the ways their frames of reference contribute to what and how they teach. Chapter 3 deals with student perceptions of the patterns of communication. It demonstrates how student perceptions of classroom activities effect both learning participation and learner performance in the classroom. Chapter 4 examines how students' culturally acquired knowledge and use of language shape their ways of talking, acting, interacting, and learning. It also explores how the language of the home and the language of the classroom can present discontinuities and problems in second language learning. Chapter 5 shows the extent to which maintaining the patterns of classroom communication can affect language learning and second language acquisition in general.

Part 2, chapters 6 - 8, uses the theoretical framework from Part 1 to examine the dynamics of communication in second language classrooms. Chapters 6 and 7 focus on teacher-student and student-student interactions and the roles these interactions play in classroom language learning and language acquisition. Chapter 8 places the framework for communication within the broader socio-cultural contexts of school-based issues and the community as a whole, and examines the effects of these factors on the second language classroom.

Part 3 (chapters 9 and 10) deals with promoting communication in second language classrooms. Chapter 9 examines how teachers can vary patterns of classroom communication to maximize learners' linguistic and interactional competencies. It also examines ways of increasing students' opportunities for language learning and language acquisition. Chapter 10, the last chapter of the book, examines how second language learners can develop their classroom communicative competence through an understanding of the nature of classroom communication.

Johnson bases her assumptions and conclusions on actual classroom discourse. Her goal is to develop more successful teaching practices by promoting more effective communication between teachers and students

in second language classrooms. For example, she stresses the need for teachers to allow for greater variability in the patterns of classroom communication (p. 90). Rather than adopt a rigid and implicit approach to teaching, teachers should make the goals and structure of the class explicit to the learners and create instructional activities that allow for greater variability in both task performance and social participation. This encourages teachers to utilize a variety of methods and techniques to accommodate different learning styles and language outcomes.

The book has many strengths. First, it demonstrates that classroom communication does indeed have a system. Even allowing for spontaneity and variability, classroom communication is not a chaotic event, though second language students may perceive it as such. Second, the classroom communication system can be described using classroom data. This can lead to a better understanding of the system and to more effective teaching strategies.

Third, the book explains why learners do not always learn what teachers teach. The implications are that this usually occurs when teachers do not address the many variables involved in classroom communication that effect learning. Using classroom-based data, Johnson shows that teachers' perceptions of the "what" and the "how" to be learned in classrooms differ from those of their students. Furthermore, the perceptions differ from student to student. One role of the teachers is to be aware of these variables and to accommodate them in their teaching processes. This is the heart of the book. Suggestions for achieving it include setting up optimal conditions for learning and language acquisition (p. 87). and making explicit the socio-cultural norms underlying learning in the target language (p. 71).

Fourth, the book sheds light on why learners often learn what teachers do not teach. Johnson emphasizes that classrooms are a dynamic language learning environment. She points out that while teachers do play an important role, they are only one source of language learning. Students also learn a lot from each other. Student-student interactions have a direct impact upon classroom learning and second language acquisition (p. 128). This is an important consideration for all language teachers. She urges teachers to maximize the benefits of student-student interactions by creating opportunities for cooperative learning through group work (p. 112).

In recognizing the role that students have in classroom language learning, Johnson's research supports the findings of learner-centered approaches to teaching (Nunan 1988). Such research and teaching approaches advocate similar benefits from student-student classroom in-

teractions for both language learning and language acquisition beyond the classroom. Researchers have found that group work fosters language acquisition by providing learners with the opportunity for productive language use and the negotiation of meaning (Nunan 1988).

Fifth, the book benefits from the many case studies from around the world. Short but specific reference is also made to Japan, for example the differences between Japanese and Anglo-American education systems and styles of learning (pp. 53-54; pp. 59-60). This directly concerns teachers in Japan and highlights the role of the socio-cultural factors that affect language learning across cultures.

A weakness of the book is its failure to recommend *how* the classroom teacher can effectively accommodate the variables that effect classroom communication. Johnson stresses the need for teachers to establish greater variability in their patterns of communication in the classroom without telling how to do this. She tells teachers what to consider in the classroom, but not how to deal with the issues that emerge. For example, Johnson provides many useful tips about the role of modeling and group work (pp. 111-114; pp. 157-158) in setting up optimal conditions for classroom learning, yet suggests few instructional activities or pedagogical techniques. Teachers might feel overwhelmed by the enormity of the task confronting them. Recognising the complexity of classroom communication is one thing, but adjusting teaching practices to deal with is another.

Another weakness of the book is its treatment of the classroom technique of modeling. Johnson stresses the importance of teachers' modeling of classroom language and classroom tasks as a way of promoting classroom communication. Yet she does not discuss the importance of modeling appropriate text-types or genres and the impact this has on effective communication (Halliday & Hasan, 1989). Research suggests the importance of providing learners with a variety of appropriate genres of language development and effective classroom communication (Martin, 1985). The role of text-type or genre language modeling is missing from the book.

Johnson provides a valuable description of classroom communication and the many patterns that effect this in second language classrooms. The book successfully combines a theoretical framework with actual classroom practice to examine a topic of interest to teacher and researchers. It raises significant issues for second language teaching, learning, and teacher training. The insights that emerge are relevant to any teacher or teacher educator interested in understanding more about the factors that influence teaching and learning in second language class-

rooms. It uses authentic data to encourage reflection on what actually takes place in the classrooms, rather than rely on intuition or anecodotal information. From it, teachers can draw conclusions based on sound theoretical evidence, which can help them develop more effective teaching practices. Anybody interested in clasroom-based research and its role in second language teaching will find this book of value.

References

Barnes, D. (1976). *From communication to curriculum*. Middlesex: Penguin.

Halliday, M.A.K., & Hasan, R. (1989). *Language, context, and text: Aspects of language in a social-semiotic perspective*. Oxford: Oxford University Press.

Martin, J.R. (1985). *Factual writing: Exploring and challenging social reality*. Geelong, Victoria: Deakin University Press.

Nunan, D. (1988). *The learner-centered curriculum*. Cambridge: Cambridge University Press.

Sinclair, J., & Coulthard R.M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils*. Oxford: Oxford University Press.

An Introduction to Spoken Interaction. Anna-Brita Stenström. London: Longman. 1994. 238 pp.

Reviewed by
Hugh Molloy
Setagaya YMCA

This textbook/reference work (p. xiii) gives a detailed description of English conversation. Though broadly based on the 1975 model of Sinclair and Coulthard, it incorporates the work of Stenström and other writers. Stenström collected her data from the London-Lund Corpus of English conversations, a collection of about half a million words taped, transcribed, prosodically analyzed, and computerized. The English Stenström analyzes was spoken by "adult, educated British English speakers" (p. xii).

The first of the book's five chapers covers some characteristics of conversations and defines some of the common terms in conversation analysis, including *backchannels*[1] (p. 5), *pause units*, *tone units*, and *information units* (pp. 7-10), and the distinction between *coherence* and *cohesion* (p. 14). She introduces *adjacency pairs* (pp. 17-18) as well as the cooperative maxims of Grice (pp. 18-19), and the effect of temporal position (pp. 20-22), tonicity and pitch direction (pp. 24-25), and context (pp. 25-29) on the function of lexical items and utterances. Stenström covers much ground in

29 pages, perhaps too much for the uninitiated. Though a useful list of further readings appears at the end of each chapter, the number and frequency of specialized terms in this first chapter could be daunting to those unfamiliar with discourse analytic works. Subsequent chapters include useful exercises, but there are none here.

The second and third chapters are the heart of the book, treating interactional structure and interactional strategies. In short, Stenström explains how conversations and their parts fit together. To do this, she uses a five-level hierarchy consisting of, from largest to smallest, the *transaction*, the *exchange*, the *turn*, the *move*, and the *act* (p. 31). Other introductory texts based on or reproducing the Sinclair and Coulthard model such as Cook (1989) do not include the *turn* level. In Stenström the *act* is what the speaker wants to communicate; this is the "smallest interactive unit." A *move* is a speaker's contribution to a particular exchange. In a question-and-answer exchange a *move* would be everything comprising the answer or the question, possibly involving in an answer a preface, an answer proper, and a qualification. An *exchange* is a set of at least two moves by two different speakers. A *transaction* comprises all the *exchanges* having to do with a single topic (p. 30). Implicit in Stenström's hierarchy is a higher level, one comprising one or more *transactions,* the "conversation."

The *turn* here is "everything the current speaker says before the next speaker takes over" (p. 30). This addition does not quite fit with the levels Stenström has adopted from Sinclair and Coulthard. To be able to tell when an *act*, a *move*, an *exchange*, or a *transaction* has occurred, observers have to understand the language in which the interaction takes place. For example, if observers are unable to tell when one topic begins and the other ends, they cannot demarcate *transactions*. However, to understand when *turns* begin and end, observers do not have to understand a word of the language in which the interaction happens. They have only to be able to tell the difference between two voices.

As a convenience for tying together *transactions*, the *turn* is useful. Certainly the *turn* is convenient as a way to organize what we see in a "transcript." Whether the *turn* is a useful device for thinking about conversations as they naturally exist, as audio or audiovisual events, is another question. When we observe a conversation in time, be it on tape, on video, or live, having in hand the concepts of *transactions*, *exchanges*, *moves*, and *acts* is useful. These allow us to organize what we experience, categorize what we hear, and make some sense of the seeming mess of conversation. Furthermore, each one is a category to which we assign one or several linguistic or paralinguistic realizations

in a conversation. Accumulating records of many such realizations in a given category can add to our knowledge of how language works. It does not seem as if the *turn* adds to our knowledge in this way. The concept of a *turn*, furthermore, is redundant when we observe conversations by listening or watching. As long as we can tell one voice from another, there is no need to have a special term for when one person talks. The *turn* is probably essential only in analyzing transcriptions of conversations.

Eleven pages in the second chapter are devoted to an inventory and examples of *act* types. Stenström notes three categories of *acts*: *primary acts*, which "can realize moves on their own," *secondary acts*, which "accompany and sometimes replace primary acts," and *complementary acts*, which "accompany but rarely replace primary acts" (pp. 38-39). The inventory comprises 28 primary *acts*, seven secondary *acts*, and 10 complementary *acts*.

Some of the categories seem more accommodating than others. A <statement>[2] "informs or expresses opinion" (p. 40). An <opine>, "gives one's personal opinion." An <evaluate> "judges the value of what the previous speaker said." A <react> "expresses attitude and strong feelings" (p. 39). Stenström gives one example of each act, not enough for me to be able to tell if, for example, <statement> and <opine> are to be considered different kinds of acts or if the latter is a special case of the former. Chapter 3 includes many more examples, however, and it is possible to skip back to Stenström's inventory with these in mind.

The reader should be aware that though some of Stenström's terms for *acts* are the same as those used in the Sinclair and Coulthard model, the referents for those terms are not the same in the two models. For instance, in Stenström, an <accept> "agrees to a <request>, <suggest>, etc" (p. 39), while an <accept> in the Sinclair and Coulthard model "shows [the teacher[3]] has heard correct information" (Cook, 1989, p. 47).

Stenström continues the second chapter with descriptions of *exchanges* and *transactions* and concludes with a discussion of discourse markers and interaction signals. She emphasizes that words she calls interaction signals, words such as "right," "quite," "really," and "OK" (p. 59) "can do different things in the same place" (p. 61). As examples, she offers instances of the use of "right" as *moves* such as a [respond], a [re-open], a [follow-up], as *acts*, such as a <confirm>, an <emphasizer>, and a <question> (pp. 61-62) and notes that it can furthermore function as an <appealer>, an <accept>, an <answer>, as well as a discourse marker (p. 62). This inventory of the functions is misleading. What seem to be multiple citations of one word are actually citations of sev-

eral different ones. The intonation of the phonemes that make up the sounds we transcribe as "right" so affect meaning that we would be better off considering differently intoned instances as separate words. As with the case of the *move*, the work done here with "right" and other interaction signals and discourse markers would not be necessary were it not for the limitations of the recording medium. Hearing "right" intoned as a request for confirmation, as a topic closer, or as an appeal for agreement, observers would have no difficulty in immediately recognizing the function.

Much of the work done with this type of discourse analysis involves repairing the damage done by a dictionary approach wherein words meanings have been cataloged simply by orthography. From Hasan's idea of instantial lexical relations (McCarthy & Carter, 1994, pp. 203-204) and the contention that "[e]very lexical item may enter into a cohesive relation . . ." (Halliday & Hasan, 1976, p. 288), it seems clear that dictionaries often ill serve us. Dictionaries that took into account factors of intonation and formulaic language would obviate perhaps half the seeming slipperiness Stenström attacks. "Obviously, it is not possible to classify the lexical items [that function as interactional signals and discourse markers] into clearcut functional categories" (p. 67). I would disagree. It is not so obvious, unless we limit ourselves to the tools available today.

Chapter 3 is devoted to a discussion of how turn-taking works: how speakers take up, hold, and yield turns (p. 68). Especially helpful are the discussions of how *exchanges* can begin (pp. 88-109). I have found it easy to understand how conversations proceed within *exchanges*, however, how conversations move from one exchange to another has not been as obvious. Stenström points out language functions besides questions which initiate *exchanges*. Her discussion helped me recognize why my students' exchange procedures sometimes seem stilted. More importantly, it assisted me in thinking of ways to help students overcome this. To go further in understanding how exchanges proceed, however, more attention to how paralinguistic features of spoken interaction—the hems and haws—work to prepare conversant for *exchange* openings and terminations, was needed.

Chapter 4, devoted to conversation, includes a discussion of strategies for introducing and terminating topics (pp. 151-154). The focus here is akin to those of writers who take a functional-notional approach to language and language teaching such as van Ek (1975), but with greater attention paid to the strategic function of language. Stenström includes some information on how paralinguistic features, in particular pitch direction, function (see especially pp. 167-168). While the curi-

ous reader could get a more complete list of mechanisms whereby turn-taking functions from other sources such as Coulthard (1985), Stenström makes her list immediately useful by providing numerous examples.

Stenström also writes on the workings of three different categories of spoken interaction: interviews, discussions, and conversations (p. 169), examining one example of each from the London-Lund corpus. Presumably, the examples selected are typical, but as I read I found myself wondering how well other interviews, discussions, and conversations would fit her analyses.

The short, 11 page, final chapter is a discussion of discourse and grammar. Stenström analyzes several extracts from conversations in terms of her discourse units, mainly acts, tone units, and clause units, providing convincing support for a clause-, rather than sentence-, based approach to the grammar of spoken interaction.

As an introduction to one approach to the analysis of spoken language, this book serves well enough, provided the reader has outside support and keeps in mind that there are other approaches and even other manifestations of the same approach. The density of information might prove intimidating even for experienced readers, but the exercises provide a chance to assimilate the information. As a reference work for those who have some knowledge of conversational analysis, *An Introduction to Spoken Interaction* serves rather as an equivalent of a birder's field guide than as a work of ornithologic reference. This is a student's guide: The suggestions for further reading should prove useful. However, with no references in the bodies of the chapters, anyone hoping to quickly find information on particular points will be disappointed. Provided one is satisfied with the one approach used, this book is useful. It is possible, however, to question the utility of that approach for those who hope to teach how to do spoken interaction rather than to talk or think about it.

Current approaches to conversation analysis, of which Stenström's book is a good example, are a help to us as language analysts. As second language teachers, however, we need more. It is not enough for us to know what has happened in conversations. We need to know what happens in conversations as they happen. We need to know not only what conversations are like, but especially what participating in a conversation as it unfolds is like, how language is used on the fly, before the conversation is finished and ready for picking apart and analysis. How to use language is what we teach, after all. Analyzing language product does not go far enough.

Notes

1. *Italicized* terms appeared in bold in Stenström.
2. Angled brackets symbolized *acts* and square brackets symbolized *moves* in Stenström.
3. The Sinclair and Coulthard model was originally drawn up from evidence collected in a number of British primary schools (Cook, 1989, p. 46).

References

Cook, G. (1989). *Discourse*. Oxford: Oxford University Press.

Coulthard, M. (1985). *An introduction to discourse analysis* (2nd ed.). London: Longman.

Halliday, M.A.K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

McCarthy, M., & Carter, R. (1994). *Language as discourse: Perspectives for language teaching*. London: Longman.

Sinclair, J., & Coulthard, R.M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils*. Oxford: Oxford University Press.

van Ek, J. A. (1975). The threshold level. In C. J. Brumfit & K. Johnson (Eds.), (1979). *The communicative approach to language teaching* (pp. 103-116). Oxford: Oxford University Press.

A Communicative Grammar of English, 2nd Edition. Geoffrey Leech & Jan Svartvik. Harlow, Essex: Longman Group Limited, 1994. 423 pp.

Reviewed by
Kevin Varden
Meiji Gakuin University

One of the most difficult things for an English learner to master is understanding the use of intonation to convey meaning. Grammar references are usually of little help, since they restrict themselves to the *syntax* of English: word order, collocations, verb agreement, among others. Enter the most useful aspect of *A Communicative Grammar of English*. This is not to categorize the work as only a reference of intonation—it by no means ignores traditional grammatical information, it simply adds another welcome dimension. (For examples of a third grammar of paralinguistic gestures, see Clark, Moran, & Burrows, 1981, pp. 253-270).

The authors have contributed to two other recent grammars, *A Student's Grammar of the English Language* (Greenbaum & Quirk, 1990) and *A Comprehensive Grammar of the English Language* (Quirk, Greenbaum, Leech & Svartvik, 1985). However, the organization and communicative focus of this grammar make it more than simply a condensed version.

Changes in the Second Edition

Users familiar with the first edition's format will have no difficulty using this edition. The separate entry format for each topic remains, although the total number of entries has been reduced from 886 to 747 and the number of pages has grown from 324 to 423.

Reorganization is evident in the Table of Contents. Instead of listing Part One only as "Varieties of English" the authors retitle it more descriptively as "A Guide to the Use of this Book" with individual listings of each entry. A shortened "Note on Phonetic Symbols" is entered here. The authors include and expand entries previously in Part Two of the first edition, including "Intonation" (a general discussion of its uses). Part Three, "Grammar in Use," has become Part Two, and Part Four, "Grammatical Compendium," has become Part Three with the more student-friendly title "A-Z in English Grammar." Of note is the deletion of the decidedly student-unfriendly use of abbreviations in the discussions of verb patterns (e.g., T1 for a transitive verb with a single object), along with the somewhat cryptic table explaining them.

Another change is the introduction of corpora data, following the trend of newer grammars such as Collins COBUILD (Sinclair, 1990). This provides the user with uncontrived examples of English usage, a welcome edition. As a result, where appropriate, examples contain the errors, hesitations, and false starts that can not be avoided in normal speech. Separate entries discussing these have been added to assist the user. However, only a portion of the examples are new; many examples remain unchanged from the first edition.

Like the first, the second edition contains cross-references to a larger grammar, in this edition to *A Comprehensive Grammar of the English Language* (Quirk, Greenbaum, Leech & Svartvik, 1985). In keeping with the times, the authors have removed the generic use of the male pronoun in examples and included a separate discussion on non-sexist usage. All in all, the grammar has been given an attractive facelift. Whether or not this justifies spending the money to replace a first edition is questionable since much of the information it contains has not changed.

Contents

The main parts of the grammar, "A Guide to the Use of this Book," "Grammar in Use," and "A-Z in English Grammar," are preceded by a "List of Symbols" used throughout and followed by the index. There are no references, recommended further readings, or corpus acknowledgments, a minor failing that could have been easily remedied.

The symbols and abbreviations used are for the most part standard; neither instructors nor students should experience any difficulty using them. The use of angled brackets to indicate the variety of English (<written>, <spoken>, <formal>, <informal>, <BrE> for British English, <AmE> for American English, etc.) may be unfamiliar, as may the use of single and double vertical bars (| and ||) to delineate tone unit boundaries and separate dialectal differences respectively. When making points about intonation, raised and lowered text help clarify the intonational contours. In addition, the cuisenaire-rod style symbols used in the discussion of tense and aspect are much more helpful than simply describing them in words.

The work contains excellent discussions on many of the techniques speakers use when communicating orally. This includes the use of many filler and repair words, a refreshing inclusion—their use may not be grammatical, but they cannot be avoided in real-world oral communication. The authors included written communication in the definition of the Communicative Method, and so the text contains many references to literary communicative techniques. They discuss literary style in a separate entry as well as in notes throughout. There are also entries on similarities and differences between the two modes of communication.

The text differentiates between British and American English via the abbreviations <BrE> and <AmE>. Other dialects of English are not represented, since ". . . the varieties of English used in the United States and in Britain are the most important in terms of population and use throughout the world . . ." (p. 28). Whether this position is defensible or not, and whether or not other dialects contain noteworthy grammatical differences, as opposed to differences in lexicon or pronunciation, I do not have the experience to say.

Part 1: "A Guide to the Use of this Book," contains information necessary to use the grammar as well as the background information needed to understand the entries. This includes discussions of the inherent differences between oral and written modes of communication; interactive, non-interactive, and cooperative uses; characteristics of spoken English; intonation; the use of phonetic symbols; geographical differ-

ences, and the uses of formality, politeness, and tactfulness. All in all, quite a useful summary of the use of the English language in itself.

Linguistically, the discussion on intonation is simplistic in the use and meanings of the three tonal patterns it distinguishes, but it is a thorough introduction nonetheless. Examples are typically marked for nuclear stress, tone, and intonational group; the use of raised and lowered text mentioned earlier aids understanding. A great deal of information is presented in a manner accessible to a Japanese college student.

Part 2 contains the main grammar section. There is a wealth of information contained in this section, called "the central part and the largest part of the grammar" by the authors (p. 6). It contains four sections in itself, and can be visualized as four concentric circles, each adding a layer of complexity to the communicative act.

Section A: "Concepts," represented by the innermost circle, contains discussions of concepts associated with our experience in the world. These include notions of nominal reference, abstractness, quantity, definiteness, relations between ideas expressed by nouns, restrictiveness, tense and aspect, duration, frequency, spatial relations, causality, condition, degree, role, comparison, inclusiveness, and topic marking. In particular, the diagrams that accompany discussions of nominal reference, quantity, degree, and spatial relations are small but clear, and quite useful.

Section B: "Information, Reality, and Belief," contains discussion of the giving and receiving of information within the context of the logic that we continually process information with. Entries here include statements, questions and responses; omission of information; reported statements and questions; denial and affirmation; agreement and disagreement; fact, hypothesis, and neutrality; degrees of likelihood; and attitudes of truth. The heavy use of stress marking in many examples accurately reflects the use of intonation in providing non-lexical information.

Section C: "Mood, Emotion, and Attitude," discusses the pragmatic use of the language in achieving a given task. This includes emotive emphasis in speech, describing emotions, volition (conscious choice), permission and obligation, influencing people, friendly communications, and vocatives (forms of address). Again, since intonation is crucial in the display and interpretation of moods, emotions and attitudes, the authors include stress-marked examples.

Finally, Section D: "Meanings in Connected Discourse, " represented by the outer-most circle, discusses the context sensitivity of sentences to surrounding material—the place of an utterance in the communication as a whole. Topics are linking signals, linking clauses and sentences, "general purpose" links, cross-reference and omission, presenting and focusing in-

formation, and order and emphasis. The entries under "cross-reference and omission" should be particularly helpful to Japanese students, since English "reference at a distance" is often problematic. An entry on reference to understood knowledge, another major source of non-understanding is missing and would improve the grammar. (Try asking your students to explain the slogan "Budweiser: It's it, and that's that.")

Part 3: "A-Z in English Grammar," resembles a more traditional grammar reference, with the addition of intonational marking on various examples. It provides the user with an alphabetically arranged discussion of many grammatical topics. There are, among many other things, an excellent discussion on the three groups of determiners (grouped by position in a noun phrase), an extensive listing of irregular verbs both alphabetically and by group, and a thorough listing of nationality words. This section also contains a large number of cross-references, to both entries within this text and in the larger *A Comprehensive Grammar of the English Language* (Quirk et al., 1985).

Several typefaces are utilized in the index: italics for individual words, all capitals for grammatical terms, plain text for functions and meanings, and angled brackets to indicate entries on variety. Thoughtfulness in preparation is evident; for example, although the authors use the term *concordance* in the text instead of *agreement*, there are entries for both in the index.

The book is very free from errors. In consulting it extensively, I only came across one cross-referencing error—the discussion of willingness (entry 320, p. 161) refers the reader back to entry 129 on uses of *will*; this entry is actually entry 140. I disagree with the continued usage of the traditional label *verb phrase* (p. 396) to refer to combinations of verbs (i.e., main verbs and their associated modals and auxiliaries) since the term is now ambiguous between this usage and the broader definition of the whole of the predicate used in generative grammar.

These minor caveats aside, this is a reference book that I both use and refer to my students. It provides much of the information students need to continue developing their communicative ability. Instructors will find it an excellent resource as well. In particular, the inclusion of intonation is a great improvement over standard grammars. I recommend this work to Japanese students over either of the two larger works the authors have contributed to; while the grammatical information contained in them is more detailed, the communicative focus of this work makes it more applicable to the needs of the average student. However, whether one would want to spend the money to replace a first edition of this work would probably depend on how worn out the covers are.

References

Clark, R.C., Moran, P.R., & Burrows, A.A. (1981). *The ESL miscellany*. Holt-Saunders International Edition. Tokyo: Holt-Saunders Japan Ltd.

Greenbaum, S., & Quirk, R. (1990). *A student's grammar of the English language*. Harlow, Essex: Longman.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1972). *A grammar of contemporary English*. Harlow, Essex: Longman Group Limited.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. Harlow, Essex: Longman.

Sinclair, J. (1990). *Collins cobuild English grammar*. London: Harper Collins Publishers.

Functional English Grammar: An Introduction for Second Language Teachers. Graham Lock. Cambridge: Cambridge University Press, 1966. 296 pp.

Reviewed by
Randall Davis
Nagoya City University

As part of the Cambridge Language Education Series on teaching, *Functional Grammar* attempts "to provide teachers with ways of thinking about English grammar" and to aid us in "understanding the kinds of problems learner may have" associated with it (p. xi). However, while Lock's aim is admirable, he fails to provide a text that is within the linguistic and practical reach of layman readers, presumably ESL teachers, which is something the title and preface clearly suggest.

The back cover entices the reader saying that the book has been written to cover "areas of difficulty for second language learners" by providing examples, tasks, and teaching discussion questions to explore these particular problems. Chapter 1, "Some Basic Concepts" defines functional grammar, outlines how grammar is used to express meaning, and defines a number of basic "basic" terms. The next eleven chapters attempt to examine grammar and how it is used, including

chapters such as "Doing and Happening I: Transitivity of Action Processes" (Chapter 4), "Doing and Happening II: Ergativity, Phrasal Verbs, and Phrase" (Chapter 5), and "Representing Time: Tense and Temporal Adjuncts" (Chapter 8). With such complex topics, the book is in fact better as a classroom text on grammatical structure for linguistics than for self-study for classroom teachers looking for an introductory guide to English grammar.

To be fair, the latter half of the book is a little easier to read and more interesting with its focus on different kinds of speech acts, expressions of attitude, opinions and judgments, and textual meaning, all as part of functional grammar. Even then, it is heavy reading, which tends to alienate the "introductory" reader and does little to dispel the criticisms against grammar teaching which Lock discusses in Chapter 13.

Ironically, the most informative and thought-provoking part is this chapter "Issues in the Learning and Teaching of Grammar." In his overview of the history of grammar teaching, Lock notes that although a structural grammar methodology "has tended to be associated with a more or less discredited approach to second language teaching," there is a resurgence of interest in the area. He goes on to point out several methodological options for teaching language, emphasizing that the teaching of grammar needs to be interwoven into the instruction of the other language skills through "meaningful and motivating activities" (p. 277) that provide opportunities for students to solve problems by creating generalizations about how grammar works. One idea that Lock refers to is an information transfer technique where students transfer spoken or written input to charts or graphs (e.g., for a presentation) and thus expose themselves to certain grammatical features.

Unfortunately, such moments are few and far between. *Functional English Grammar* turns out to be a reference tool appropriate for the researcher or academic linguistic, not the introductory guide to grammar teaching teachers want to help them class. The promises contained in the subtitle and on the covers, that readers need "no previous study of English grammar or linguistics" are unmet as readers wade through chapter after chapter trying to decipher the complex explanations of "functional" grammar. To end on a bright note: readers looking for a truly "functional" book on grammar can find several listed in the references that do fit the bill.

Spoken English on Computer: Transcription, Mark-up and Application. Edited by Geoffrey Leech, Greg Myers, and Jenny Thomas. Essex, England: Longman Group Limited, 1995. 260 pages.

Reviewed by
Charles Adamson
Miyagi University

Corpora, large bodies of language data (often annotated) stored in computer readable form, are becoming increasingly important. Linguists are using them to investigate language as it actually is used, rather than as they think it might be. At conferences and workshops, there are frequent references to corpora. Recently textbook writers and classroom teachers have begun deciding what should be taught on the basis of data from corpora. It appears that this trend will continue and a knowledge of corpora and their uses will be a necessary part of a TEFL/TESL teacher's education and hence a matter of interest for JALT members.

Spoken English on Computer (SEOC) is a good introduction to the field and a must read for anyone working with any of the established corpora of spoken language or considering the development of a new one. The book began as a collection of papers that were presented at a *Workshop on Computerized Spoken Discourse* held in England in September, 1993. The editors, however, did not feel that the papers covered the field sufficiently so they elicited papers from other workers in the field. SEOC is the result.

Each of the book's three sections has an introduction by the editors: Part A, with seven articles addressing fundamental issues; Part B, with seven articles discussing the application of corpora of spoken language; and Part C, with six articles describing corpora which exist or are being constructed. While the subject matter is quite technical, the authors make the content accessible to non-specialists.

The first four chapters of Part A analyze the theory and principles of transcription. The most problematic area is what data should be included in the corpus and what form that data should take. This applies especially to para-linguistic information. The next two chapters deal with the recommendations of the Text Encoding Initiative (TEI) and their implications. The TEI is the result of joint development by a number of groups and outlines a markup method based on Standard Generalized Mark-up Language (SGML). While this treatment is insufficient to allow the reader to actually mark up texts, it is an excellent introduction and would be of great value to any teacher or researcher using a corpus that uses TEI. In the final chapter of the section, John Sinclair discusses eco-

nomical ways to construct corpora of spoken language, the need for better definitions of "spoken material" and the need to make corpora more user-friendly.

Part B balances the theory from Part A with practical examples of the use of spoken language corpora including the study of language impairment, code-switching, and the structure of conversation. There are also discussions of the tagging of time, intonation, and word class. Another chapter describes a research project in which conversations were recorded while native speakers were doing information gap activities on maps. The editors point out that these chapters when taken together raise a number of practical problems: the quality of the original recordings; confidentiality for the sources; the importance of information that is lost through transcription, coding, and mark-up; and the hardware that is required.

Part C contains descriptions of six corpora of spoken language: (1) The London-Lund Corpus; (2) The COBUILD spoken corpus; (3) The Machine Readable Spoken English Corpus (MARSEC); (4)The International Corpus of English; (5) The BNC spoken corpus; and (6) The Bergen Corpus of London Teenager Language (COLT).

More than 500 references have been gathered together in a Bibliographical References section at the end of the book along with an Author Index and a Subject Index. The latter, however, is not particularly useful since it has only about 120 entries.

In the General Introduction the editors indicated that various strands of research are currently coming together in the field of spoken language corpora. This book, which provides chapters discussing some of the important aspects of many of these strands, is an excellent overview or introduction to the field for language teachers and researchers in Japan.

Books to Review

Reviews are being sought of the following books. Contact the Reviews Editor (see Guidelines) for further information.

Baynham, M. (1995). *Literary practices: Investigating literacy in social contexts.*

Brown, J.D. (1996). *Testing in language programs.*

Durand, J., & Katamba, F. (1995). *Frontier of phonology: Atoms, structures, derivations.*

Freeman, D., & Richards, J.C. (1996). *Teacher learning in language teaching.*

Grabe, W., & Kaplan, R.B. (1996). *Theory and practice of writing.*

Gutknecht, C., & Röle, L.J. (1996). *Translating by factors.*

Hasan, R., & Williams, G. (Eds.). (1996). *Literacy in society.*

Hatch, E., & Brown, C. (1995). *Vocabulary, semantics, and language education.*

Jacobs, R.A. (1995). *English syntax: A grammar for English language professionals.*

Kunnan, A.J. (1995). *Test taker characteristics and test performance: A structural modeling approach.*

McKay, S.L., & Hornberger, N.H. (eds.) (1996). *Sociolinguistics and language teaching.*

Nunan, D., & Lamb, C. (1996). *The self-directed teacher: Managing the learning process.*

Pennington, M.C. (1996). *Phonology in English language teaching: An international approach.*

Quirk, R. (1995). *Grammatical and lexical variance in English.*

Ridout, R.M. (ed.) (1996). *The Newbury Hourse dictionary of American English: An essential reference for learners of American English and culture.*

Sarangi, S., & Slembrouck, S. (1996). *Language, bureaucracy and social control.*

Sasaki, M. (1996). *Second language proficiency, foreign language aptitude, and intelligence: Quantitative and qualitative analyses.*

Spada, N., & Frölich, M. (1995). *COLT (Communicative orientation of language teaching) observation scheme: Coding, conventions and applications.*

Street, B.V. (1995). *Social literacies: Critical approaches to literacy in development, ethnography and education.*

Swan, M. (1995). *Practical English usage (2nd ed.).*

Thomas, J., & Short, T. (Eds.). (1996). *Using corpora for language research.*

van Lier, L. (1996). *Interaction in the language curriculum: Awareness, autonomy and authenticity.*

Wodak, R. (1996). *Disorders of discourse.*

Information for Contributors

Contributions must conform to the *JALT Journal* Editorial Policy and Guidelines.

Editorial Policy

JALT Journal, the research journal of the Japan Association for Language Teaching (*Zenkoku Gogaku Kyoiku Gakkai*), welcomes practical and theoretical articles concerned with foreign language teaching and learning in Japanese, Asian, and international contexts. It welcomes well written contributions which provide readers with a link between theoretical and practical issues, especially those addressing current concerns in pedagogy, methods, and applied linguistics. Areas of specific interest are:

1. curriculum and teaching methods
2. classroom centered research
3. cross-cultural studies
4. teacher training
5. language learning and acquisition
6. overviews of research and practice in related fields

The editors encourage submission of full-length articles, short research reports, book and media reviews, essays on trends in language education, reports of pedagogical techniques which are thoroughly framed in theory and supported by descriptive or empirical data, and comments on previously published *JALT Journal* articles. Occasionally, *JALT Journal* will issue a Call for Papers for theme-based issues. Articles should be written with a general audience of language educators in mind, with statistical techniques and unfamiliar terms clearly explained or defined.

Guidelines

Style

JALT Journal uses the *Publication Manual of the American Psychological Association* (available from the Order Department, A.P.A., 1200 17th St. N.W., Washington DC). Consult recent copies of *JALT Journal* or *TESOL Quarterly* for examples of documentation and references. Give the page numbers of cited works in both the text and references.

Format

No more than 20 pages (approximately 6000 words) for full-length articles, including reference list, typed and double-spaced on A4 or 8.5"x11" paper. *Research Forum* submissions should be no more than 10 pages. *Perspectives* submissions should be no more than 15 pages. Authors must supply camera-ready diagrams or figures (if any) before final publication. The author's name and references that identify the author should appear only on the cover sheet.

Materials to be submitted

Three (3) copies of the manuscript
Cover sheet with title, running head title (one to four words), and author name(s)
Abstract (no more than 150 words)
Japanese translation of title and abstract if possible (less than 400-*ji*)
Biographical sketch(es) (no more than 50 words)
Authors of accepted manuscripts must supply a disk copy (RTF or ASCII).

Evaluation Procedures

All manuscripts are first reviewed by the editorial staff to insure they comply with *JALT Journal* guidelines. Those considered for publication are subject to blind review by at least two readers, with special attention given to the *JALT Journal* Editorial Policy, the significance and originality of the submission, and the use of appropriate research design and methodology. Evaluation is usually completed within three months.

Restrictions

Papers submitted to *JALT Journal* must not have been previously published, nor should they be under consideration for publication elsewhere. *JALT Journal* has First World Publication Rights, as defined by International Copyright Conventions, for all manuscripts published. We regret that manuscripts or computer disks cannot be returned.

Address for Manuscripts and Inquiries

Tamara Swenson, *JALT Journal* Editor
Osaka Jogakuin Junior College
2-26-54 Tamatsukuri, Chuo-ku, Osaka 540 Japan

Submissions to *Perspectives*

The editors invite reports of pedagogical techniques which are thoroughly framed in theory and supported by descriptive or empirical data to the *Perspectives* section. Send submissions to:

Sandra Fotos, *JALT Journal* Associate Editor
Senshu University, 2-1-1 Higashi Mita
Tama-ku, Kawasaki, Kanagawa 214 Japan

Reviews and Address for Reviews

The editors invite reviews of books, tests, teaching systems, and other substantial publications in the field of language education. A list of publications which have been sent to JALT for review is published monthly in *The Language Teacher*. Please send requests for books, materials, reviews guidelines, and queries to submit reviews to:

Thomas Hardy, *JALT Journal* Reviews Editor
Tamagawa University, 6-1-1 Tamagawa Gakuen
Machida-shi, Tokyo 194 Japan

Japanese-Language Manuscripts

The *JALT Journal* welcomes contributions on teaching Japanese as a second language (JSL). Submissions must conform to the above Guidelines. Authors are requested to proved a detailed abstract in English, 500-750 words in length. Refer to the Japanese-language Guidelines for details. Send all Japanese-language manuscripts to:

Naoko Aoki, *JALT Journal* Japanese-language Editor
Shizuoka University, Department of Education
836 Oya, Shizuoka 422 Japan

Address for Inquiries about Subscriptions or Advertising

JALT Central Office
Urban Edge Bldg. 5F
1-37-9 Taito, Taito-ku, Tokyo 110 Japan
Tel: 03-3837-1630; Fax: 03-3837-1631
(From outside Japan: Tel: 81-3-3837-1630; Fax: 81-3-3837-1631)

JALT Journal 日本語論文投稿要領

JALT Journal は、日本語教育に関する日本語論文の投稿を受け付けます。第二言語教育／学習研究の文脈を踏まえ、しっかりした研究計画に基づいている、実践への応用の可能性を示した理論的・実証的研究や理論的裏づけを持った実践報告などで、教育、教授法、応用言語学などの今日的問題を扱ったものを歓迎します。

文体は一般的な学術論文のスタイルを使ってください。章だてのしかたや参考文献のデータの書き方などは、Publication Manual of the American Psychological Association (4th. ed.) の定める方式にできるだけ近い形にしてください。ご不明の場合は、JALT Journal の英語論文を参考にするか、日本語編集者までお問い合わせください。また、JALT Journal の読者は現場の教師が主ですから、特殊な専門用語や統計的手法は、わかりやすいように定義したり説明したりしてください。

原稿の長さは、参考文献リストも含め400字づめ原稿用紙30枚以内です。A4の用紙に横書きで、1行40字、1ページ30行で印刷してください。手書きの原稿は受け付けられません。図表を入れる場合は、JALT Journal のページのサイズに合わせて作成してください(縮小コピー可)。執筆者の名前や所属機関名は、原稿には書かないでください。ご提出いただくものは、以下の通りです。

- 原稿3部
- 執筆者の名前と所属機関名を書いた表紙
- 400字以内の和文要旨
- 英文のタイトルと、500～750語の英文要旨
  英文要旨が用意できない場合は、400字づめ原稿用紙3枚程度の詳細な和文要旨をご提出ください。
- 100字以内の執筆者略歴
- 可能ならば、Macintosh のテキスト形式で保存したファイル

査読の方法は、編集委員会で投稿要領にあっているかどうかを確認したあと、少なくとも二人の査読者が査読を行います。査読者には執筆者の名前は知らされません。査読の過程では特に、原稿がＪＡＬＴ Journal の目的にかなっているか、言語教育にとって意味があるか、独創性はあるか、研究計画と方法論が適切かが評価されます。査読は通常二か月以内に終了します。

JALT Journal に投稿する原稿は、すでに出版されているものや他の学術雑誌に投稿中のものは避けてください。JALT Journal は、そこに掲載されるすべての論文に関して国際著作権協定による世界初版権をもちます。

お送りいただいた原稿は返却できません。

投稿原稿の送り先とお問い合わせは以下にお願いします。

〒422　静岡市大谷836　静岡大学教育学部
JALT Journal 日本語編集者　青木直子

JALT Journal 第19巻　第1号

1997年5月　6日　印刷

1997年5月15日　発行

編集人　タマラ・スウェンソン
発行人　ジーン・ヴァン・トロイヤー
発行所　全国語学教育学会事務局
〒110東京都台東区台東1－37－9アーバンエジビル5F
TEL (03)3837-1630; FAX (03)3837-1631
印刷所　コーシンシャ株式会社
〒530 大阪市北区天満2-13-3 TEL (06)351-8795

The 23rd Annual
International Conference
on Language Teaching /
Learning & Educational
Materials Exposition

JALT 97

Oct. 9 - 12

Act City
Hamamatsu, Japan

Theme:
Trends & Transitions

FOR MORE INFO CONTACT:
JALT Central Office
Urban Edge Bldg. 5th Fl
1-37-9 Taito, Taito-ku, Tokyo 110
Tel: 03-3837-1630 FAX: 03-3837-1631

Visit the JALT97 WWW Homepage at:
http://www.miyazaki-mic.ac.jp/JALT/JALT97.html