

ISSN 0287-2420

¥750

JALT Office

# JALT JOURNAL

JOURNAL OF THE JAPAN ASSOCIATION  
OF LANGUAGE TEACHERS

全国語学教育学会

VOLUME 12 NO.1

MAY 1990

昭和五十四年四月二十一日 第三種郵便物認可  
第十二卷 第一号 平成二年五月一日発行  
THE LANGUAGE TEACHER 特別増刊号

**CONTENTS**

Editorial ..... 5

**Articles**

- Teaching Presentation and Discussion Skills to EAP Students  
*Daniel Horowitz and Lynn Stein* ..... 9
- Process, Literature, and Academic Realities:  
The Contribution of Daniel Horowitz  
*Ann M. Johns* ..... 29
- Challenging the Task  
*Liz Hamp-Lyons* ..... 37
- Reading in a Foreign Language: Research and Pedagogy  
*Patricia L. Carrell* ..... 53
- Japanese-English Code Switching in Bilingual Children  
*Sandra S. Fotos* ..... 75
- Pausology and Listening Comprehension:  
Theory, Research, and Practice  
*Roger Griffiths* ..... 99
- Where Do Tests Fit into Language Programs?  
*James Dean Brown* ..... 121

**Book Reviews**

- Language Transfer: Cross-Linguistic Influence in  
Language Learning (Terence Odlin)  
*Reviewed by Fred E. Anderson* ..... 141
- Bilinguality and Bilingualism (Josiane F. Hamers  
and Michel H.A. Blanc)  
*Reviewed by Mary Goebel Noguchi* ..... 146
- ESP in the Classroom: Practice and Evaluation  
(Dick Chamberlain and Robert J. Baumgardner, eds.)  
*Reviewed by Sarah Rilling* ..... 151

**Information for Contributors.** ..... 154

**Editors**

Malcolm Benson  
*Hiroshima Shudo University*

Charles Wordell  
*Nanzan University*

**Book Review Editor**

Jane Wieman

**Japanese Language Editor**

Kyoko Nozaki

**Editorial Board**

Charles E. Adamson, Jr. <i>Trident College, Nagoya</i>	Richard Berwick <i>Momoyama Gakuin, Osaka</i>
John Fanselow <i>Columbia University</i>	John Fry <i>British Council, Kyoto</i>
Michael Horne <i>Nagoya University</i>	Ann Johns <i>San Diego State University</i>
Susan Johnston <i>Temple University Japan</i>	Teruyuki Kume <i>Kobe University of Foreign Studies</i>
Michael Liggett <i>University of Tsukuba</i>	John Maher <i>International Christian University</i>
Peter McCagg <i>International Christian University</i>	Joy Reid <i>University of Wyoming</i>
Lawrence Schourup <i>Kobe University</i>	

Publications Chairperson, Ann Chenoweth  
*The Language Teacher* Editors,  
Eloise Pearson, Ann Chenoweth

# NATIONAL OFFICERS 1990

## President

Deborah Foreman-Takano

## Vice President

Tatsuya Komatsu

## Treasurer

Philip Crompton

## Recording Secretary

Rita Silver

## Programs

Linda Viswat

## Membership

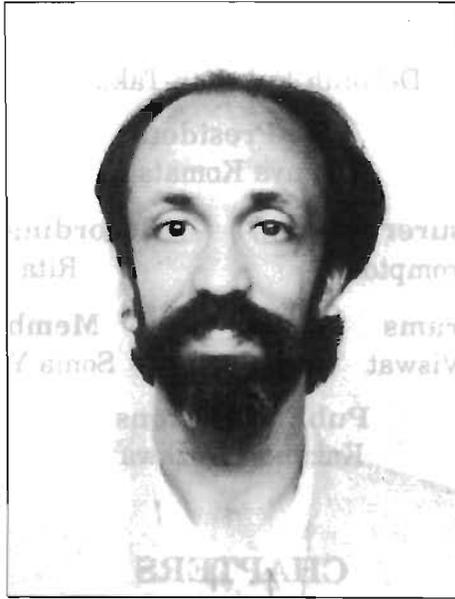
Sonia Yoshitake

## Public Relations

Kazunori Nozawa

## CHAPTERS

Chiba, Fukui  
Fukuoka, Gunma  
Hamamatsu, Himeji  
Hiroshima, Ibaraki  
Kagoshima, Kanazawa  
Kobe, Kyoto  
Matsuyama, Morioka  
Nagano, Nagasaki  
Nagoya, Nara  
Niigata, Okayama,  
Okinawa, Omiya  
Osaka, Sapporo  
Sendai, Shizuoka  
Suwa, Takamatsu  
Tokushima, Tokyo  
Toyohashi, Utsunomiya  
West Tokyo, Yamagata,  
Yamaguchi, Yokohama



## **In Memory of Daniel M. Horowitz**

**1950—1989**

Danny Horowitz, co-editor of the *JALT Journal* from May of last year, died of heart failure on October 22, 1989. He is survived by his wife Takako and son Aaron. Danny had a brief, splendid career as a teacher, writer, and editor, and is missed deeply by the many people who loved him. This issue includes an article he co-authored and two articles written in his honor.

## **Editorial**

This issue is dedicated to the memory of **Daniel Horowitz**, *JALT Journal* co-editor for just one issue before his untimely death. We—all of us who knew him as colleague, scholar, and friend—mourn his passing.

At the time of his death Danny was working on a number of projects with faculty members of International Christian University. One such project involved a new departure for Danny, whose previous work had largely concerned aspects of ESL writing. In “Teaching Presentation and Discussion Skills to EAP Students,” **Daniel Horowitz** and **Lynn Stein** present a carefully thought out procedure by which students, even reticent ones, become involved in making short presentations and contributing to a discussion in the foreign language. Underlying the ideas in the paper is the belief—which many will share—that EAP students need academic speaking skills; further, curriculums that fail to develop these areas are failing to prepare students for full academic participation.

In a summative article on Danny’s work, **Ann Johns** pays tribute to his acumen in the field of ESL writing, pointing out in considerable detail the extent of his contribution. In the hectic and competitive world of academic writing, the product is always more important than the process. ESL students, Danny argued, are entitled to a training which enables them to perform within the academic paradigm. Such students, Danny felt, were vulnerable in the area of writing, both by the formulaic nature of the demands made on them, and by the unforgiving nature of the academic environment. His empathy with the students’ viewpoint is what set him apart from many who think about ESL academic writing.

In another tribute to Danny’s involvement in ESL writing, **Liz Hamp-Lyons** describes continuing research into essay examination prompts. Drawing on her own earlier research, she explores how certain students respond to a prompt with a “challenge” by questioning the validity of the prompt. A “challenge” response (as opposed to an “unmarked” one) poses problems for the reader/grader, possibly resulting in overmarked or undermarked scripts.

Of particular interest to language teachers is **Patricia Carrell's** ongoing research into second language reading, in particular schema theory. "Reading in a Foreign Language: Research and Pedagogy," first heard at JALT '89, summarizes many of the stages through which the research of the past twenty years has passed. Such has been the nature and relevance of the ideas produced during that period that virtually every modern ESL/EFL reading text shows their influence. However, teachers who routinely use techniques such as pre- and post-reading activities may be unaware of the complexity of the work that established them. This paper, therefore, provides a useful resource for those already familiar with the main ideas of schema theory, and a comprehensive account of it for those who are not.

A phenomenon frequently encountered in Japan, but not well understood, is code switching, a bilingual's alternation between two languages. In a preliminary study of four bilingual (Japanese-English) children, **Sandra Fotos** shows how, far from being a random activity, switching is used by the children to achieve functional ends. Further, a number of the established theories concerning switching are found to be inadequate in explaining what occurs across languages that are as syntactically different as Japanese and English.

Consciously or unconsciously, everyone who speaks knows that speech rate affects comprehension. Language teachers are aware of this too, but trainee teachers get little specific guidance on how to speak in the classroom, apart from general admonitions to speak slowly and be clear with instructions. In "Pausology and Listening Comprehension: Theory, Research, and Practice," **Roger Griffiths** reviews the literature on speech rates in considerable detail, and provides us with an update on research in this very basic area. Potentially, such research will be of great value to teachers of listening comprehension per se, to examination bodies, and to language teachers generally, in their day-to-day interaction with language learners.

With the complexity of the demands made on administrators, teachers, and students, nobody wants to waste time, or to have their time wasted. Underlying **James Dean Brown's** article,

"Where Do Tests Fit Into Language Programs?" is the idea that one aspect of being efficient in English language teaching is to test appropriately. But appropriate testing demands a knowledge of the many test types and an understanding of the various purposes for which testing takes place. Brown here continues his energetic campaign to educate language teachers in these areas.

This issue also contains reviews by **Fred Anderson**, **Mary Goebel Noguchi**, and **Sarah Rilling**, treating books on language transfer, bilingualism, and English for special purposes, respectively.

The editors thank Hisako Okuda and Steven Ross for their help in the preparation of this issue.

### Correction

There were two misprints in the article by Richard Berwick and Steven Ross, "Motivation After Matriculation: Are Japanese Learners of English Still Alive After Exam Hell?" The last line of page 200 is ming. The two full sentences should read:

In the case of structure, the strongest predictor was a general interest in English as a foreign language. Of the small amount of variance accounted for by the survey items, the desire to study overseas was the strongest predictor of listening proficiency.

In Table 6, page 204, the heading "Structure" should be over the two left columns, and the heading "Listening" should be over the two right columns.

## CHOOSE THE BEST



### BBC BEGINNERS' ENGLISH

- Two stages at false beginners' level
- Listening and speaking carefully integrated

### THE LOST SECRET

- Unforgettable English in a gripping two-hour video
- Students' book available in bilingual and monolingual editions
- Comprehensive teacher's guide



To: **MEYNARD PUBLISHING LIMITED**

Yanagiya Bldg. 6F, 2-18-6 Takadanobaba, Shinjuku-ku, Tokyo 160  
Tel: 03-232-3611 Fax: 03-232-3805

*Please send more information;*

BBC BEGINNERS' ENGLISH  THE LOST SECRET

**Name:** \_\_\_\_\_

**School:** \_\_\_\_\_

**Address:** \_\_\_\_\_  
(School/Home)

## **Teaching Presentation and Discussion Skills to EAP Students**

**Daniel Horowitz and Lynn Stein**  
*International Christian University*

This paper argues that academic speaking and discussion skills should be included in the EAP curriculum and describes an oral presentation activity that gives students practice in these skills. Working from the premise that such skills, though complex, are amenable to practice and improvement, the authors prepared a handout that gives students detailed instructions for preparing and giving a presentation and in leading a discussion, and designed a peer evaluation form for both formative and summative feedback. Based on student evaluations of the activity, it is concluded that it provided a useful and exciting language learning experience.

### **1. Introduction**

Although “getting them talking” is one of the major goals of lower level speaking/listening classes, teachers of more advanced classes—especially English for Academic Purposes (EAP) classes—may wonder about the importance of speaking skills for their students. On the minus side, Johns’ (1981) and Ostler’s (1981) surveys of essential academic skills consistently found speaking skills to be less important than listening, reading, or writing. However, a careful examination of Ostler’s data reveals that although oral skills—giving talks, participating in panel discussions, discussing issues in class, and asking questions in class (p. 495)—were considered important only by about one-third of the undergraduates surveyed, these figures rose sharply for graduate students. Partial confirmation of Ostler’s data is provided by Westerfield (1989), who found that “MBA professors require active participation in the oral treatment of a case,” and that oral participation may determine “up to 25% of an MBA student’s course grade” (p. 75). And since there is “a growing number of international graduate students and a declining pool of undergraduates” (Kaplan 1989, p. 110), it is clear that academic oral

skills should be taught in any program that has the goal of preparing students for university work.

This article describes an activity that helps students practice essential academic speaking skills and which, as their own evaluations attest, gives them that feeling of excitement and accomplishment so essential to successful language learning. The activity is called *Presentation and Discussion*, and it consists of each member of a group of three to five students giving a 5-minute presentation and then leading the group in a 10-minute discussion. Although there is nothing remarkable or original about this basic idea, we believe that four things account for our success with it: first, sufficient time is spent teaching students what is expected of them; second, students are given opportunities for practice so that they can fulfill those expectations; third, students receive useful feedback from their peers; and fourth, the activity is repeated every few weeks so that students can use the feedback they received to improve their performance over time.

The activity has three phases: *preparation*, *presentation*, and *feedback*. Both the preparation phase and the feedback phase make use of handouts, which are described in the first part of this article. Classroom procedures for the day of the activity are discussed in the second part and students' reactions in the last section.

## 2. Preparation

Sitting through a series of poorly prepared oral presentations is a truly stultifying experience for teachers and students alike, and this may explain why many teachers are simply not willing to risk class time for this type of activity. A more positive choice, however, lies in recognizing that giving a good presentation and leading an effective discussion involve complicated skills, that these skills must be learned and practiced, and that mastering them requires time and well-organized instruction. Thus, the handout students receive (Appendix), which may seem excessively long and detailed, is in our experience welcomed by students. Since class time is devoted to discussing it, students overcome their initial fear of its length and are grateful for the explicit guidance it contains.

The remainder of this section consists of a discussion of questions teachers may have about the handout. We suggest that

readers read the handout (Appendix) first and then come back to this section. The questions are grouped under headings that correspond to the sections of the handout.

### *2.1 Introduction of the Activity*

#### *How should the handout be used?*

We give students the handout to read for homework, asking them to highlight or underline the parts that seem most important to them and to mark any places they do not understand. Since the handout is quite long, we try to spread the discussion of it over two or even three class periods. This initial investment in class time pays the most benefits when the activity is used several times per term, because most students need only to review the handout briefly for subsequent presentations.

#### *Why do students work in small groups rather than speak to the whole class?*

There are both theoretical and practical reasons for preferring small group to whole class work. Long and Porter (1985) present a number of arguments supporting the idea that group work aids second language acquisition, noting that students get more practice with the language, use a greater variety of language, and engage in more negotiation for meaning (pp. 221-222). Bejarano (1987) also found that

peer discourse in the small-group setting apparently permitted pupils to think aloud and, by necessity, to talk in less polished language. There was far less inhibition and tension under these circumstances because discourse served communicative needs rather than the demands of public recitation. (p. 495)

Finally, practical reasons: small-group work allows an entire class to finish the activity in one class period; also, it limits the negative effects of uninteresting, repetitive, or inadequately prepared presentations to the fewest possible victims.

#### *What determines the exact size of the group?*

The exact size should be determined by the length of the class period. We can fit four fifteen-minute blocks in our 70-minute classes, so we divide students into groups of three and four, with as many groups of four as possible. This allows 10 minutes of "administrative" time.

*Why is the fifteen minutes divided into a 5-minute presentation and a 10-minute discussion period?*

This seems to strike a reasonable balance between speaking and listening. By the end of the class, each student has spent about 5 minutes in the “hot seat,” 10 minutes as a discussion leader, 15 minutes listening to others’ presentations, and 30 minutes taking part in discussions.

## *2.2 Choosing a Topic*

*How much freedom do students have in choosing their topics?*

With the one (admittedly significant) restriction that their topic must in some way relate to the content topic they are presently studying, they have almost complete freedom. These presentations are part of the writing and reading/discussion classes at International Christian University, which focus each term on three content topics (one term’s fare: English as an international language, artificial intelligence, and bioethics). These topics provide the basis for all reading, writing, and discussion assignments. As a counterbalance to academic writing assignments, in which there is often no place for students’ personal experiences, we encourage them to use personal experience or knowledge as the core of their presentations on each topic in the reading/discussion classes.

We have not found it necessary to “approve” topics beforehand, but on some occasions we have asked students to tell us their topic the day before the presentations so that we could make sure that the students in each group had different topics.

## *2.3 Preparing Your Talk*

*What guidance are students given in organizing their talks?*

Given the wide variation of topics that students choose to speak about, there is no specific pattern that can be recommended to all. We therefore fall back on the concept of “say what you are going to say, say it, and say what you said.”

*Why are students required to use an index card?*

When our Japanese students think of oral presentations, they think of speech contests, where speakers compete in giving excruciatingly emotional and painstakingly rehearsed formal speeches. One of the main goals of the activity described here is to help our students learn a less formal alternative, and the use of an index

card with speaking notes only (no complete sentences) fits in with this goal. Students must learn to use the card only as a prompter, not a script, and we have learned to watch for the student who brings a beautifully written and well-organized note card on the day of the presentation and then proceeds to read from a fully written-out script! Thus, we warn students against either memorizing or writing out their talk, and the index card provides just enough room for a happy medium.

*Why five discussion questions?*

We have found that five questions provide sufficient fuel for ten minutes of discussion. It may be that five is too many, but preparing fewer than five discussion questions minimizes the potential breadth of discussion.

#### *2.4 Practicing Your Talk*

*Why are students required to submit a practice tape?*

Since the success of this activity is largely due to the practice that students do before the presentation, we want to make sure that their practice is as effective as possible, and to us this means practice *out loud in English*. We explain that silent practice does not simulate performance conditions closely enough, whereas practice out loud with a tape recorder gives them the opportunity to listen critically to themselves, just as they will be critically listened to by their peers. The practice tape also gives them a chance to time themselves.

Practicing in front of a mirror is a technique that one of the authors learned from studying to be a professional singer. It has two benefits. First, it allows students to see themselves as others will see them. Second, it helps students practice making eye contact—one of the most difficult aspects of Western body language to teach to Japanese students—in a non-threatening environment.

#### *2.5 During the Talk*

*What problems do students face as speakers and listeners?*

Unfortunately, our students tend to identify classroom “speaker—listener” situations with the “all knowing and powerful teacher—passive pupil” situations they were exposed to in high school. Thus, we find it necessary to re-educate students about their roles as speakers and listeners, preparing them for the idea that it is

normal and desirable for speakers and listeners to *interact*: that interruptions and questions are a normal part of group discussion, and that, in an academic setting, eye contact with interlocutors is the polite way to show interest in what others are saying (see Tannen, 1984, for a discussion of this and other aspects of cross-cultural pragmatics).

### 2.6 *During the Discussion*

*What problems do students face as discussion leaders and discussants?*

Most of our students tell us that in high school they *never* had group discussions in their classes. Thus, students need to be taught that a skillful discussion leader facilitates discussion rather than dominates it; exhibits flexibility in directing the course of a discussion rather than slavishly following a “plan”; looks for opportunities to involve each member of the group in the discussion in a variety of ways, rather than in a mechanical fashion by directing questions first to student #1, then student #2, and finally to student #3 ad nauseam; responds positively to opportunities for interesting and unexpected detours; and appreciates the fact that sometimes s/he can be most effective by simply allowing the discussion to proceed unimpeded.

But it is not enough to teach discussion leaders about equality and reciprocity; it is also necessary to teach discussants to value themselves as members of the audience—that each presentation and discussion is for *their* benefit, not for the (sole) benefit of the leader. When students finally come to realize that they have the right to direct the discussion toward aspects of the topic that interest them and to speak among themselves, they feel freer to say what is on their minds, to question each other, to probe more deeply than their culture normally permits in “polite” conversation.

### 2.7 *Feedback and Grading*

*What kind of feedback do students get? How are they graded?*

The simple answer to this question is that students get feedback from each other, both because we want to emphasize that the students’ primary responsibility is to their audience of peers, and because it is physically impossible for us to monitor all the student groups at the same time. The form they use in evaluating each

member of their group has places for both formative (the left side of the form) as well as summative (the right side) evaluations. The former are collated and returned to the student without the evaluator's name, while the latter are seen only by us.

This somewhat elaborate scheme is our accommodation to our students' culturally-induced reticence to judge each other critically. The built-in anonymity helps us "sell" our students on the idea that formative evaluation is essential for improvement and that summative evaluation is a necessary part of our educational situation. Of course, teachers who do not need to give a grade might want to re-design the form, perhaps using only the formative section.

Students receive three grades, one as *presenter*, one as *discussion leader*, and one as *discussant*. All three grades are based on their peers' summative evaluations, and the first two are also influenced by our judgment of their level of preparation as indicated by their index cards. Although the quality of their practice tape does not affect their grade, failure to submit a tape (or submitting a blank one!) lowers their *speaker* and *discussion leader* grades.

### 3. Classroom Procedures for the Day of the Activity

*What can teachers do before the presentations begin?*

First, teachers may want to assign students to groups. For example, if students have submitted their topics before the day of presentation, teachers can form groups of students with a good mix of topics. Another reason for controlling group membership is to make sure that over the course of a term, students speak to a variety of classmates, not just their best friends. Finally, in a multi-lingual classroom, teachers may want to form groups of students from a variety of language backgrounds.

On the day of the presentation, teachers should set up desks (or chairs and tables) *with an extra chair in each group* so that they can easily join each group during the activity. We do this from the beginning of the term so that students become accustomed to our being part of their audience.

*How can the teacher keep things running "on schedule"?*

Each presentation/discussion should begin and end at the teacher's signal. In practice, this means that the teacher or the

group must decide on the order of presenters; the teacher must signal the beginning of each presentation; the presenter must use a watch to make sure that the presentation stays within its time limit; and the teacher must give a "two-minute warning" (after about 13 minutes) so that discussion leaders can begin their summaries. After 15 minutes, the teacher signals the end of the first time block, and the process begins again after a momentary break.

When some groups have one fewer member than the rest, the smaller groups will finish one time block before the larger ones. When this happens, members of the smaller groups should redistribute themselves among the larger groups for the final time block. However, students should evaluate only the members of their original group.

*What does the teacher do during the activity?*

If the class is not too large, teachers can circulate quickly at the beginning of each time block and hear at least a portion of each student's presentation. They can spend a little more time in each group during the discussion period and may even want to contribute to the discussion, although we do so only occasionally for fear of being perceived as the (temporary) "leader." While listening, teachers might want to jot down some notes to give to individual students to help them with problem areas not covered by the peer evaluation form.

Teachers can also use the very short break between presentations to point out good things they have just seen or heard, such as a discussion leader who showed the flexibility to allow a discussion to "stray" in an interesting way, or a discussant who challenged something another discussant said.

*How is the "paperwork" handled?*

At the end of the class period, students put their practice tape and note card(s) into the envelope they were instructed to bring to class. When the envelopes have been collected, we distribute evaluation forms (the same form as in the handout, only on a separate sheet of paper), and students write the name of each group member on a separate form. Students complete the forms as homework for the next class.

#### 4. Student's Reactions

For all that one may say about the relevance of an EAP activity to students' future academic needs, one should not let the product of learning overshadow the process. As Hutchison and Waters (1987) point out, students "should get satisfaction from the actual experience of learning, not just from the prospect of eventually using what they have learnt" (p. 48). In this section, we present two types of evidence, a questionnaire and a number of comments written by students, both of which suggest that students do indeed value this activity as a satisfying learning experience.

Table 1 shows the responses of 43 students to five questions about the activity on an end-of-term questionnaire.

Clearly, the great majority of students found the activity "useful" and felt that it had helped them improve their "speaking" and "discussion" skills. However, there were fairly large groups of students who felt it had not helped them improve their "listening" skills and who were not happy about the grading of the activity. The feelings behind both these responses are revealed by comments made by students in their weekly written class evaluations.

##### 4.1 *Negative comments*

The fact that a number of students did not feel that their listening ability had improved may be related to the varying degrees of ability among students. As one pointed out,

I couldn't improve my listening skill because some of us have poor pronunciations while the others have good ones.

Another possible explanation is that they spent their time listening to each other rather than to "real" speakers of English. This problem may be avoided if the teacher emphasizes that much of the usefulness of learning English comes from its importance as an international medium of communication, and that, therefore, interaction with non-native speakers is every bit as valuable as that with native speakers.

Several other students commented on their feelings about being in groups with more advanced students. One claimed that the differences were "more than what can be solved by efforts." Another student said:

**Table 1**  
Student Responses to End-of-Term Questionnaire

	Agree	Agree strongly	Disagree	Disagree strongly
Presentation/Discussion was a useful activity.	23/53%	17/40%	2/5%	1/2%
Presentation/Discussion helped me improve my speaking skills.	14/33%	21/49%	6/14%	2/5%
Presentation/Discussion helped me improve my listening skills.	1/2%	19/44%	21/49%	2/5%
Presentation/Discussion helped me improve my discussion skills.	10/23%	28/65%	4/9%	1/2%
The grading of Presentation/ Discussion was fair.	1/2%	25/58%	16/37%	1/2%

Of course I have something to talk about, but my poor ability and envy of (very fluent students) make me stop talking!

The second major area of negative feelings concerned the *subjectiveness* of the grading process:

The grading of presentation/discussion were sometimes not fair because different person have different way of judgment and depends person in group we get different grade....

I was sometimes afraid of classifying other's ability. ..It would be better to add other evaluation (more objectively).

This response may be peculiar to Japanese students, who have passed through an educational system where *all* evaluation has at least the facade of objectivity, or it may be a more universal feeling

that, as “mere” students, they are not yet capable of evaluating their peers.

Although one way to eliminate this complaint would be for a teacher to grade each presentation from an audio tape, we feel that both the teachers’ time and the dynamics of the discussion group would be sacrificed, and that little in the way of “objectivity” would be gained. It seems more reasonable to impress upon students the fact that their teachers’ judgments are no more objective than theirs, that they are better off being judged by three careful listeners than by one teacher, and that their best guarantee of fairness is to fill out each others’ evaluation forms as conscientiously as possible. This is a complaint that will probably not go away, but if it is dealt with forthrightly, it can lead to fruitful discussion and learning about the importance of speaking as *communication* rather than *recitation*.

#### 4.2 Positive Comments

Positive comments about the activity far outnumbered negative ones, and only a small sample can be given here. Students noted their improvement:

In this term, presentation is just done second time, but I see big changes. What is most important is our attitude to try to speak and express ideas in English as much as possible.

They noted changes in their understanding of cultural aspects of communication:

I was a little surprised that it is considered POLITE behavior to stop a speaker to ask clarifying questions. Because it is considered polite in Japan to be *just* a listener when the other person is speaking. So I asked a lot of questions to become a POLITE person in English way.

They expressed their and their classmates’ excitement:

Today’s class was very exciting. Each presentator had a keen interest and showed it as well as possible.... These three presentation really enriched me....

They wrote about their interest and involvement:

We discussed enthusiastically, almost forgetting we were talk-

ing in English. Isn't that great?

This week I was afraid of presentation/discussion. But I could spent good time during the class. Because each of our group didn't speak fluently but could speak eagerly and ask question, give answers friendly.

They approved of practicing with a tape recorder:

Because I had to record in the tape, I had to practice many times and as I practiced I became to know or understand what I was going to talk about exactly.

When I listened to my speaking from a tape, I could easily find my poor pronunciation.

And they claimed that the activity stimulated them intellectually:

When I planned about my talking, I thought about the article more deeply than usual, and I could notice many interesting point....

Discussion was very helpful to make my own idea more clearly.

## 5. Conclusion

This activity clearly requires a serious investment of time and energy by both teacher and students, but it is time and energy well spent. One student's words seem to sum it up best:

The most useful and successful aspect of this class for me is presentation/discussion. Thanks to that, I became able to speak or discuss in front of other people without hesitation or being shy. I now want to do that again....

## Acknowledgements

*The authors would like to express their thanks to Susan Kocher and Minako Sakata for reading and commenting on earlier drafts of this paper.*

**Daniel Horowitz** taught at International Christian University in Tokyo. He was co-editor for the *JALT Journal* and served on the

editorial board of the *English for Specific Purposes Journal*. His articles on academic reading and writing appeared in those journals as well as the *TESOL Quarterly* and *System*.

**Lynn Stein** teaches in the English Language Program at International Christian University in Tokyo. She holds a Master's in TESL from San Francisco State University and has presented papers at various conferences throughout Asia.

### References

- Bejarano, Y. (1987). A cooperative small-group methodology in the language classroom. *TESOL Quarterly*, 21, 483-504.
- Hutchison, T., and Waters, A. (1987). *English for Specific Purposes*. Cambridge: Cambridge University Press.
- Johns, A.M. (1981). Necessary English: A faculty survey. *TESOL Quarterly*, 15, 51-57.
- Kaplan, R. (1989). The life and times of ITA programs. *English for Specific Purposes* 8, 109-124.
- Long, M.H., and Porter, P.A. (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quarterly*, 19, 207-228.
- Ostler, S.E. (1980). A survey of academic needs for advanced ESL. *TESOL Quarterly*, 14, 489-502.
- Tannen, D. (1985). The pragmatics of cross-cultural communication. *Applied Linguistics*, 5, 189-195.
- Westerfield, K. (1989). Improved linguistic fluency with case studies and a video method. *English for Specific Purposes*, 8, 75-83.

## Appendix

### Presentation and Discussion

- I. Introduction
- II. Choosing Your Topic
- III. Preparing Your Talk
- IV. Practicing Your Talk
- V. During The Presentation
- VI. During The Discussion
- VII. Feedback And Grading
- VIII. Things to Bring on the Day of the Talk

#### I. Introduction

This term you'll each be giving several 15-minute talks related to the topics you're studying. The fifteen minutes will be divided into two parts, the **presentation**, and the **discussion**.

1. The **presentation** lasts about five minutes. During this time you (the *Speaker*) present your ideas to a small group of classmates (the *Listeners*), speaking IN ENGLISH from notes you've prepared on one or more index cards.

2. The **discussion** lasts about ten minutes. During this time you (the *Discussion Leader*) lead your group (the *Discussants*) in a discussion of your presentation, based on questions you've prepared. You conclude your talk with a brief summary of the main points that were made during the discussion.

You'll also **evaluate** each member of your group. These evaluations will help you improve your next presentation and will help me give you a grade.

The purpose of this handout is to give you step-by-step instructions for this activity. Please read it carefully!

#### II. Choosing Your Topic

You may choose any topic connected with the topic you're studying and the related readings; however, your talk should NOT be a summary of those readings. It SHOULD be about something that you're really interested in, something that **you and only you** can contribute to everyone's understanding of the topic. In choosing, ask yourself if there is anything you can actually **show** your classmates, such as a photograph or a drawing you've made. Using a visual aid such as this will make your presentation even more interesting. Finally, remember that you have a time limit of five minutes and that you must speak only English, so choose a topic which you

can comfortably present. You might, for example, talk about:

- \* something in your personal experience that helps to explain one or more of the ideas connected to the topic.

- \* your *interpretation of/reaction to/disagreement with* one or more of the ideas in the readings.

### III. Preparing Your Talk

#### 1. Do ALL of your preparation **in English**.

2. Review the readings on the topic. Note the main ideas. Try to find one or two ideas that have some personal meaning for you, that you reacted to strongly, or that remind you of something you've experienced. Then, list your ideas, just brainstorming for interesting things to say. After that, organize the best of those ideas into an *introduction, body, and conclusion*:

- \* In the **introduction**, briefly state the main idea of your talk in one or two sentences, and mention the topics you'll discuss.

- \* The **body** should be the main part of your talk. It should be organized so that your listeners can easily follow and understand your ideas. This is where you'll explain your ideas and opinions, give your examples, and show any visual aids you've brought. In choosing examples, be sure to choose ones that are easy for your listeners to understand.

- \* The **conclusion** is your last opportunity to tell your listeners what you think were the important points of your presentation. Ask yourself, "What do I want them to remember?" It's a good idea to summarize briefly the main points of your talk and emphasize, through what you say (the words) and how you say it (your tone of voice and your gestures), those points which were most important to you. By doing so, it will be clear to your listeners that your presentation is over and that it's time to begin the discussion period.

3. Write your presentation in note/outline form on one or more medium-size index cards (look at the sample note card below to get an idea of what yours should look like). Your NAME, SECTION, and presentation TITLE should be written across the top of the first card. Also, the three parts of your talk—INTRODUCTION, BODY, and CONCLUSION—should be clearly labeled on the left side of the card. Draw a line between these labels and your notes so that it will be easy to see them as you speak. Feel free to use both sides of the card and/or additional cards.

It's very important to remember that you will NOT be reading your cards during your talk—you'll only be looking at them to find your place and your next idea, so you should:

\* Leave lots of empty space. It's very difficult to find your place when your card is crowded with words.

\* Leave enough room **between** sections to make additions at any point.

\* Mark the main points with stars or arrows or big black circles so that you can keep track of where you are.

\* Write your notes in **dark pencil**, in **large letters**. It's easy to make changes if you've written in pencil and large letters are easy to see. Write only to the *right* of the line.

**Sample note card  
for "Nature and Humanity" presentation**

NAME:	SECTION:	TITLE:
<u>INTRO:</u>	The nature around us is made by humans—it's not real	
<u>BODY:</u>	3 parts to talk—nature in home town, daily life, on earth	
	**In my home town Example: experiences with firefly	
	** In our daily life—power without humanity	
	** On the earth—environmental destruction, pollution	
	Examples: + seals' death + acid rain + amazon desert + abnormal climate	
<u>CONC:</u>	**Humans think can control nature—wrong! need harmony with nature and good relationship	

4. Write (at least) **five** discussion questions following your speaking notes on the index card. Each question should encourage your classmates to share their opinions with you and the other members of the group. Here are some suggestions for writing good ones:

\* Try to make each question *stimulating* (something that people want to talk about) and *specific* (something that people have **definite ideas** about).

\* Vary the form of the questions. Begin with different phrases such as "What do you think about...?"; "Why...?"; "Do you...?"; "Where...?"; "When...?"; "How...?"; "Have you ever...?"

\* If you ask "yes-no" questions, be ready to ask a follow-up question. Look at Discussion Question #1 on the card below for an example of a

follow-up question.

### Sample discussion questions for “Nature and Humanity” presentation

#### DISCUSSION QUESTIONS:

- (1) Did you have experiences with real nature as a child?  
Yes: give answer  
No: experience from television? books?
- (2) What do you think about the environment around you?  
Is it man-made or real nature?
- (3) Do you feel a cultural difference between Europe and Japan in treating nature?
- (4) How can we get along with nature?
- (5) What is your opinion about environmental pollution?

#### IV. Practicing Your Talk

Once you’ve written your note cards, you need to practice **what** you’re going to say and **how** you’re going to say it. For this step, you should practice with a tape recorder. Find a quiet place, set up your recorder, turn it on, take out your note cards, and start practicing. **Practicing** does NOT mean making a “perfect” recording or reading your cards. It DOES mean talking, stopping, hesitating, making mistakes, searching for a better word, changing your ideas (and maybe your notes).

Practicing also means learning how to use your note cards. Practice looking down at your card for your idea and then looking up at your classmates, BEFORE you start talking. Finally, be sure to practice each part of your talk, **including** your conclusion and your first discussion question. When you’re finished, listen to the tape. Did you speak loud enough? Will others understand you? Is the organization of the talk clear?

When you’ve practiced enough to feel comfortable with what you want to say, practice once (or twice) more with the tape recorder, index cards, and any visual aids you will use, but this time, practice in front of a mirror. Choose three or four objects (books? pencils?) to represent your classmates and put them around you as though you were in a discussion group. Then put your watch in front of you. The purpose of practicing this way is to:

\***See yourself** as your classmates will see you. Is your head buried in your index card or are you really just **referring** to it?

\* **Practice looking** at your classmates. Did you look at each book or pencil about the same number of times?

\* **Practice using** any visual aids you’re going to show. As you hold up

each one, make sure that EACH person sees it, but don't cover your face with it.

\* **Practicing timing** yourself. Remember, *you're* responsible for limiting your presentation to about five minutes.

At the end of this practice session, listen to the tape and evaluate yourself as a Presenter, using the left side of the Presentation/Discussion Evaluation Form on page 5 of this handout. If you hear a problem, try to correct it. And finally, make sure that all the notes on your index card(s) are written in dark pencil and are written large enough so that you can easily see them.

## V. During the Presentation

When you're the *Speaker*, remember to use your index card notes only for reference, to put your watch on your desk and time yourself, and to *look* directly *at* and *speak* directly *to* each of your classmates. If a group member interrupts you to ask a question, answer it as well as you can, and then continue your talk.

When you're a *Listener*, put your own index card(s) and other materials where you can't see them, *look at the speaker*, and give him/her your total attention. If you don't understand something that is said, ask for an explanation. It's much better to ask for clarification than to sit silently, pretending to understand.

## VI. During the Discussion

When you're the *Discussion Leader*, your responsibility is to get the **other members** of your group to talk about the ideas you presented, so you should probably talk the **least**. Your prepared discussion questions are the **basis** for the discussion, but you should be flexible and willing to "go with the flow" of the discussion. For example, if a group member asks an interesting question about a related topic, let the discussion go in that direction until you feel it's time to return to your original prepared questions.

During the discussion it's still your responsibility to keep track of the time. Whether or not you finish discussing your five discussion questions, be sure to leave about two minutes at the end of your presentation to *briefly* summarize the main points that were made during your discussion. Your final summary is the conclusion to your presentation and without it, your presentation is incomplete. I'll be giving you a "two-minute" warning to help you remember when to begin the summary. After the summary, you might end your presentation with a simple "thank you."

Presentation/Discussion Evaluation Form

<b>Presentation/Discussion Evaluation Form</b>	
Your name: _____ Section: _____ family name first name	
Student's Name: _____ family name first name	Check the one best description in <i>each</i> of the three sections below:
These areas need improvement:	NOTE: This side will not be shown to the speaker!
1. As a PRESENTER <i>Content:</i> _____ unoriginal topic ___ unclear introduction ___ too abstract ___ unclear conclusion ___ not well organized  <i>Technique:</i> _____ many hesitations ___ lack of eye contact ___ voice too soft ___ poor use of note cards ___ spoke too fast	This PRESENTER gave ___ an exciting, original presentation. ___ a very good, effective presentation. ___ an average presentation. ___ an ineffective presentation.
2. As a DISCUSSION LEADER ___ spoke too much ___ questions not stimulating ___ too few questions prepared ___ did not involve everyone ___ did not summarize the discussion well ___ did not "go with the flow"	The DISCUSSION was ___ stimulating and enlightening. ___ enjoyable and interesting. ___ somewhat interesting. ___ rather boring.
3. As a DISCUSSANT ___ did not participate actively ___ spoke only to the leader	As a DISCUSSANT, s/he ___ participated actively and made thoughtful comments throughout the discussions. ___ showed involvement and interest most of the time. ___ showed involvement and interest some of the time. ___ showed little involvement and interest.
4. What was the best aspect of this presentation?	_____

If you should finish discussing your prepared questions **before** the discussion period is over, ask members of your group to share more examples of personal experiences, or return to a question that was particularly interesting and ask for more opinions.

When you're a *Discussant*, show your interest by talking to **each member** of the group, not just the leader. **Be a leader** yourself by asking any questions you think of. If you're asked a question, answer it as fully and completely as you can, trying to open up new possibilities for discussion. Be willing to disagree with or challenge your classmates' ideas. Interest-

ing and stimulating discussions are those in which you ask questions, express your opinions, and disagree, as well as agree with other members of your group.

## VII. Feedback and Grading

Your homework assignment for the class following presentation day is to complete a Presentation/Discussion Evaluation Form (shown after section VIII) for **each member** of your group, excluding yourself. Although there's a place for your name on the top, it will be cut off **before** the forms are distributed to your group members.

You'll receive three grades, one as *presenter*, one as *discussion leader*, and one as *discussant*. All three will be based on your group members' evaluations of your presentation. It's extremely important that you *fairly* and *completely* evaluate each of your group members' presentations, in the same way that you expect them to evaluate yours.

In addition to the individual evaluations, your grades as presenter and discussion leader will be influenced by my judgment of how well you prepared your index card(s). I'll be looking at your notes (were they brief and clear?) and at the quality and quantity of your discussion questions (were they stimulating and interesting? Were there at least five?) Your practice tape will NOT affect your grade, unless you don't submit one! You'll receive each of your classmates' evaluations within a few days of the presentation, as well as your final letter grades.

## VIII. Things to Bring on the Day of the Talk

Be sure to bring your **index card(s)**, your **practice tape**, and an **envelope** to put them in. Please **write your name and section** on all three.

**Process, Literature, and Academic Realities:  
The Contribution of Daniel Horowitz**

**Ann M. Johns**

*San Diego State University*

This issue of the *JALT Journal* honors Dan Horowitz, and rightfully so. He accomplished a great deal while in Japan, and was already making his mark with this journal when he died.

However, since I do not live in Japan, I knew Dan in a broader context: in the world of academic writing and reading, in the world of tasks and professors' demands. In that international world he was also known and respected. And it is his contributions to the international academic community, and his influence on my own work, which are central to the discussion that follows.

I first met Dan at a publisher's dinner at the 1984 TESOL Conference. I can't remember the publisher's name, but I vividly recall sitting across from Dan, discussing with some pride the paper I planned to present the following morning, in which I was to deal with issues of teaching coherence and intertextuality in writing classes. Dan listened with genuine interest, then spoke of his own curricular innovations. As he spoke, it soon became clear to me that it was Dan, not I, who should be discussing coherence and the teaching of writing the next day; and, much to his surprise, I asked him to stand up during my talk and to present his approach, which, incidentally, is part of his book that was in preparation last year.

Since that time, I have always looked to Dan for research ideas and teaching approaches which I couldn't find elsewhere. Peter McCagg has spoken of him as "a brilliant person—the most clear-headed thinker I have ever had the honor to be associated with" (McCagg, 1989). I would like to add that he was also one of the most creative; he saw a real classroom need and concentrated on it until he came up with ideas which were both ingenious and workable. I never spoke with Dan about teaching without taking notes; I never heard him present formally without saying to myself, "Those are wonderful ideas!"

The greater academic ESL/EFL community first became aware of Dan in March, 1986, when he published a classic, often-quoted "Forum" piece in the *TESOL Quarterly*, where he took on the Writing Process Establishment, with humor, insight, and common sense (1986b). He later commented, "I wrote...this article out of anger at the arrogance I had encountered among the process advocates at TESOL '85 and at the blind 'bandwagon' mentality that periodically warps the critical judgment of so many members of our profession" (1986e, p. 796). In this "Forum" article, Dan argued that the writing process advocates had become as short-sighted and territorial as the Current-Traditionalists whom they replaced. If they were more clearheaded, he argued, they would recognize that the process approach is inadequate for students of academic writing since:

1. It fails to prepare students for at least one essential type of academic writing: the essay examination;
2. It fails for some types of writers, for example, those who always work from outlines;
3. It fails students who do not have certain types of writing strategies;
4. It fails to give students an accurate picture of university writing. (1986b, 141-144)

Dan concluded by warning that "teachers should be extremely cautious about embracing an overall approach which, in its attempt to develop their students' writing skills, creates a classroom situation that bears little resemblance to the situations in which those skills will eventually be exercised" (1986b, p. 144).

Responses to Dan's "Forum" article were immediate: two, by JoAnne Liebman-Kleine and Liz Hamp-Lyons, appeared in the December, 1986, *TESOL Quarterly*. Dan answered these responses with comments which expanded upon his original argument, maintaining that the process advocates are obsessed with the cognitive relationship between the writer and his or her internal world; therefore they provide no clear perspective on the "social nature of writing" (1986d, p. 788) so important to university success. His conclusion was one in which he urged open-mindedness, insisting that neither the academic view nor the process view of writing "subsumes the other or can stand alone" (1986e, p. 797).

These exchanges in *TESOL Quarterly* about composition theory and teaching set the stage for two of Dan's articles on topics central to our understanding of authentic academic writing: one, on genuine academic tasks in university classrooms (1986c), and another, dealing with the demands of the most common task for undergraduates, the essay examination (Horowitz, 1986a; also see Keller-Cohen & Wolfe, 1987). In his article on academic tasks, which appeared in *TESOL Quarterly* 20 (3), 1986, Dan noted the limitations of previous questionnaire-based needs assessment, and advocated a more qualitative, in depth approach to determining academic tasks. He suggested that by collecting and analyzing real tasks assigned in academic classrooms, an English for Academic Purposes (EAP) teacher could understand more clearly important issues such as audience, topic, task definition, and grading. Dan's qualitative technique was central to this study, but his insights garnered from the data are also enlightening. Two important conclusions which Dan made in this article have been repeated frequently in the literature. He wrote in the first:

1. Generally speaking, the academic writer's task is not to create personal meaning, but to find, organize and present data according to fairly explicit instructions. (1986c, p. 455)

By making this statement, Dan was asserting that genuine academic tasks are antithetical to the assignments and examinations in most writing classes: whereas prompts in writing classes often encourage student opinion and narrative, the tasks in students' academic classes require a much more objective view, purged of personal meaning.

A related finding was the following:

2. The most striking feature of the sample [of real tasks] was... the controlled nature of much of the writing called for. (1986c, p. 452)

Dan found, as others have, that whereas writing teachers tend to encourage as much creativity in terms of content and form as possible, instructors in academic disciplines often do the opposite. In many academic classes, Dan found, the thesis statement is provided, the audience is clear, the questions or topics to be covered are specified, the sources of the propositional material are obvious,

and there are a number of lexical constraints.

These findings, which contradict much of what can be found in published academic writing texts, have influenced a number of researchers in our discipline. In my own research, which attempts to discover students' difficulties with language and task representation as they confront new academic disciplines, I have found considerable evidence for the difficult transitions from the writing class to academic classes, to which Dan alludes.

For example, when my ethnically-diverse freshman students first entered the university, thoroughly grounded in the mandated process approach of our high schools, they predicted that their professors would ask for their opinions on essay examinations. The realities were quite different. One student, after her first essay examination, wrote in her journal that she had learned a few things about her new academic audiences:

I have learned to distance myself when writing. [This course]...has taught me to write more about facts instead of my opinion, which I used to. (Johns, in press)

Another student, accustomed to open-ended responses that encouraged student creativity of form and flexibility in dealing with arguments, reevaluated his examination in this way:

After looking over my [test] paper, I discovered that I need to spend more time in analyzing the question being asked. When I write papers, I tend to touch around the question, without answering it directly. This is an extremely bad quality, for on an exam the teachers going to want to focus on the main idea of that paper. Not just the ideas that touch around that main theme. (Johns, in press.)

These examples from my students' journals indicate that Dan's arguments should be considered seriously: only through actually interacting with academic tasks can we, and our students, understand their nature and demands.

The last of Dan's manuscripts of which I am aware was on John Swale's desk when Dan died. It will appear in *English for Specific Purposes* 9(2), a publication on whose editorial board Dan served for several years. I had seen this paper in one of its earlier

incarnations and had remarked to Dan that it was another of his “lightning rod” pieces which would stir up controversy and ruffle feathers, since it examines one of this year’s popular themes, the teaching of literature in academic classes, for what it implies and excludes.

The argument in his paper “Fiction and Non-Fiction in the ESL/EFL Classroom: Does the Difference Make a Difference?” is that “the literature debate is really a microcosm of the discourse communities debate”; it is a debate about whether academic English teachers should be involved in assisting students in understanding writing in their academic disciplines (i.e., “training their students”) or whether they should develop students’ empathy and maturity through the teaching of literature (i.e., “educating”).

The first position that Dan takes in the article is reminiscent of his earlier discussion of the process approach: the advocacy of the teaching of literature to “educate” students may be another attempt on the part of teachers to avoid facing up to their responsibilities. He suggests that instead of teaching literature, EAP teachers should assist students in understanding the specific demands of their academic classes by working with them in ferreting out the “secrets of the trade” in various disciplines (Swales & Horowitz, 1988). He acknowledges that the role of co-ethnographer in new academic cultures is a difficult one because teachers might “find themselves in the uncomfortable position of being less knowledgeable than their students” (Spack 1988, p. 37). However, Dan argues that “this is a position we are going to have to get used to—and I question the logic which says that since few ESL teachers are qualified to teach writing in the disciplines, we are therefore under no obligation to help our students to learn it themselves.”

The remainder of the article takes up other arguments made by those who advocate the teaching of literature as central to the teaching of academic English. One argument, that literature requires more of the reader in “making sense” and “negotiating meaning” than does expository prose, is dealt with in a discussion of studies of non-literary work in which Dan makes it clear that all interpretation of text requires “making of meaning,” and that much of this meaning-making requires understanding of the underlying

structure of a text and of the conventions which the text follows. He concludes the “meaning-making” argument by stating that “all discourse is a mixture of convention and creativity; all texts require rich interpretation; those readers who lack either the linguistic skills or the necessary background information will come away with impoverished interpretations.”

Another claim made by the literature advocates is that study of fiction provides an excellent way to connect reading and writing. Though no one would argue with the importance of this connection, Dan points out that the way one reads—and writes—literature is very different from the “guerrilla warfare” fashion in which we undertake—and teach—the reading of expository prose. In fiction, we read from start to finish: the schemata we develop are story grammars for literature. In expository prose, we skip around, looking for signposts throughout the text, for graphs, for charts, and for the metalanguage which leads the reader through the discourse. In short, the way we read, and what we read for, differ greatly between literature and most of the expository prose our students will find to be central to their academic success.

In this *English for Specific Purposes* article, as in his previous ones, Dan exhorts teachers to avoid unthinking acceptance of what is currently in vogue and, instead, “think hard about how the actual activities performed in class fit in with their students’ needs and wants.” He suggests that teachers ask such questions as:

1. Are the students studying English as a second or foreign language?
2. Are they presently enrolled in university classes?
3. Are they undergraduates or graduates?
4. Are they learning English for professional or vocational reasons?
5. Do they have a strict time limit on their university studies?

Dan Horowitz’s article ends, as did his others, with a plea to teachers:

If we keep (these) questions in mind when we think about using literature in ESL/EFL classrooms, our profession—ever in search of the answer—may be able to avoid another in the series of wild swings of opinion that seem to plague us. It will be a sign

of our growing maturity as a profession when new ideas—or old ones that have come calling again—are met not only with open arms but with a critical eye as well. (Horowitz, in press)

What has Dan left with us that we must heed? He asked us not to accept, but to investigate; not to surrender to what is easy, but to grow. He has inspired us to begin with classroom realities, and to use our common sense and dedication to teaching to create a classroom appropriate to student needs.

Many of my academic English colleagues and I are grateful for having had the opportunity to talk with, to read, and to learn from Dan Horowitz. Those who opposed his views respected him and dealt carefully with his reasoned arguments; those who agreed with him were encouraged and stimulated by his work. Whatever our theoretical and applied convictions, Dan's spirit of caution, skepticism, and intelligent enquiry will remain among us.

**Ann Johns** teaches at San Diego State University. She is co-editor of *ESP Journal* and has published in *TESOL Quarterly*, *Language Learning*, and *Journal of Basic Writing*, primarily concerning English for Academic Purposes.

### References

- Hamp-Lyons, L. (1986). No new lamps for old yet, please. *TESOL Quarterly*, 20, 791-765
- Horowitz, D.M. (1986a). Essay examination prompts and the teaching of academic writing. *English for Specific Purposes*, 5, 197-120.
- Horowitz, D.M. (1986b). Process, not product: Less than meets the eye. *TESOL Quarterly*, 20, 141-144.
- Horowitz, D.M. (1986c). What professors actually require: Academic tasks for the ESL classroom. *TESOL Quarterly*, 20, 445-462.
- Horowitz, D.M. (1986d). The author responds to Liebman-Kleine. *TESOL Quarterly*, 20, 788-790.
- Horowitz, D.M. (1986e). The author responds to Hamp-Lyons. *TESOL Quarterly*, 20, 796-797.
- Horowitz, D.M. (in press). Fiction and non-fiction in the ESL/EFL classroom: Does the difference make a difference? *English for Specific Purposes* .
- Johns, A.M. (in press). Coherence as a cultural phenomenon: Employing ethnographic principles in the academic milieu. In U. Connor & A.M. Johns (Eds.), *Coherence in writing: Research and pedagogical perspectives*. Washington, DC: TESOL.

THE CONTRIBUTION OF DANIEL HOROWITZ

- Keller-Cohen, D., & Wolfe, A. (1987). *Extended writing in the college of literature, science and the arts: Report on a faculty survey*. Unpublished report. University of Michigan, English Composition Board. Ann Arbor, MI.
- Liebman-Kleine, J. (1986). In defense of teaching process in ESL composition. *TESOL Quarterly*, 20, 783-787.
- McCagg, P. (1989). Unpublished eulogy to Daniel M. Horowitz. Tokyo, Japan: International Christian University.
- Spack, R. (1988). Initiating ESL students into the academic discourse community: How far should we go? *TESOL Quarterly*, 22, 29-51.
- Swales, J., & Horowitz, D.M. (1988, March). *Genre-based approaches to ESL and ESP materials*. Paper presented at the 22nd Annual TESOL Convention, Chicago, IL.

## Challenging the Task

**Liz Hamp-Lyons**  
*The University of Michigan*

Dan Horowitz was well-known for his research into essay examination prompts, and was greatly respected for the intellectual clarity of his work and for his humanistic grounding of that work in the central questions facing practitioners in ESL classrooms in colleges and universities. In this paper I review some of the work that has been done on prompt effects in ESL writing at the college or college preparatory level, focusing on just one small aspect in an attempt to move our work in this area toward a better general understanding. While I do not make explicit reference to Dan's work in the text, the collegial dialogue we maintained is an important underpinning of the paper.

There has been a great deal of research into the question of whether topic types, topics, the linguistic structure of questions, and an array of what we may group together as "prompt effects" have a significant effect on the measured writing ability of native writers of English. While far less work has been done on those same issues in relation to the writing of nonnative English users, it seems likely that the effects of prompts, if such exist, will only be exacerbated when we look at nonnative writers rather than native writers.

An overview of the field (Hamp-Lyons, in press) suggests that in first language writing assessment the trend is to treat topics, and even topic types, as equivalent. In the major ESL/EFL writing assessment programs the same trend emerges (Hamp-Lyons, op. cit.). For example, the TOEFL Program uses two topic types at different administrations of its Test of Written English, although analyses of trial prompts of the two topic types in the TOEFL Test of Written English (TWE) (Carlson, Bridgeman, Camp, & Waanders, 1985) showed that they behaved rather differently (they correlated at around .70). Reid (1989) looked more closely than Carlson et al. at the four writing prompts in the TWE study, and found that the students' tests varied significantly from topic type to topic type, even when the differences did not result in differing

score patterns. The variation was most marked for strong writers. Perhaps weaker writers have less language flexibility, while those with higher scores seem able to adapt to different topics. In looking at the issue of whether there is a significant effect on student essay test writing from the prompts used, then, it would seem that the answer depends on one's research orientation and the questions one asks as much as it does on "hard" numbers.

I can illustrate this from two studies carried out in the University of Michigan Testing and Certification Division. Both studies looked at prompts on the MELAB (Michigan English Language Assessment Battery, a test for nonnative users of English filling the same function as the TOEFL, but including a composition as a basic component rather than as an occasional, optional, extra). In the first, Spaan (1989) describes an experimental study of two prompts from the MELAB, chosen because they appeared on the surface to be dramatically different. Spaan provides a linguistic analysis to show the linguistic, cognitive, and schematic differences of the prompts, and interprets her score data as suggesting that even these prompts in fact yield scores which are significantly related.

In the second study, Hamp-Lyons and Prochnow (1990), in a post-hoc study, looked at all sets of MELAB prompts used in the period 1986-89, including those Spaan had used. They found that expert judges and student writers felt able to recognize easy and difficult topics, and that their judgments of the relative ease/difficulty of prompts were generally confirmed by score levels on prompts. While they confirm Spaan's assessment of her two topics as radically different in difficulty, considered a-contextually, by looking too at a general language proficiency measure they are able to suggest some reasons for essay scores that are less different than predicted. It seems that nonnative writers taking the MELAB writing test component assess which prompt is easier, and that students with weaker language proficiency choose the easier prompt while students with stronger language proficiency choose the harder prompt. Hypothesizing too that reader accommodation plays its part in pushing disparate prompts toward parity of treatment, they suggest that both weaker and stronger writers regress toward the mean in their writing score, relative to their scores on other language components. These two research studies

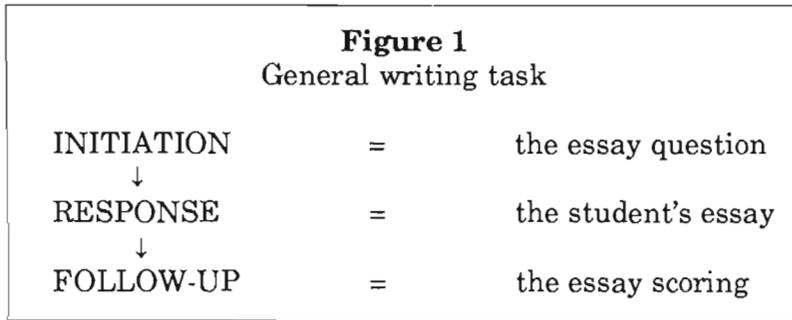
from the same testing program differ in many ways because of researchers' orientations, but only the complexity and intangibility of the kind of data we are working with when we struggle to pin down prompt difficulty permit us to arrive at such differing conclusions.

### Challenging the Task

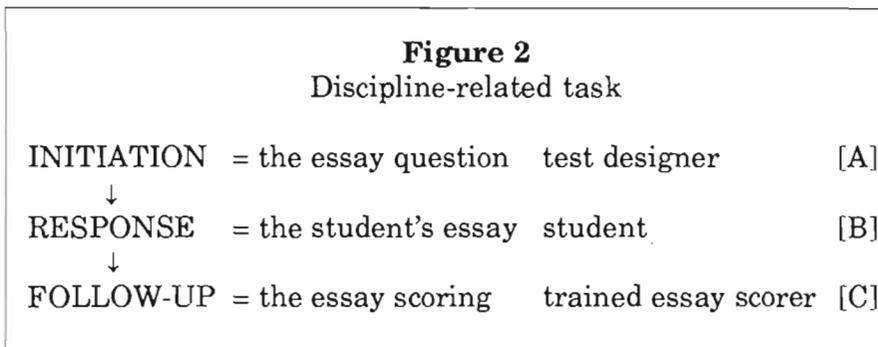
The interaction between tasks and individual writers' affective responses is so complex that it is unlikely we will ever untangle it. But to the best of my knowledge there has been **no** research on how a student's perception of herself as a member of a disciplinary community affects her response to a required task in the peculiar context of a discipline-linked writing assessment. As a very small part of a much larger study of an international writing assessment taken by nonnative English users wishing to pursue graduate study in English, I looked at this question, not because I originally planned to, but because it intruded itself upon my awareness. In so doing, I think I have arrived at some interesting and useful generalizations about the problems we face in understanding what we do when we "prompt" someone to write on a formal test.

The assessment I studied, the British English Language Testing Service examination (ELTS), aims to be fairer to students than general tests providing a variety of tasks depending on the disciplinary area the test candidate is familiar with (for a full discussion see Hamp-Lyons, 1987). The assumption is that students, particularly at the graduate level, will write better when they can write about what they know than when they must write about subjects unrelated to their studies. To investigate the claims for the test, I looked at whether test candidates got higher scores on tasks targeted on their areas of disciplinary knowledge, broadly defined, than on general writing tasks. My quantitative data yielded ambiguous results (Hamp-Lyons, 1986), and as I attempted to interpret them, I was led to look more closely at, among other things, what the test candidates had actually written.

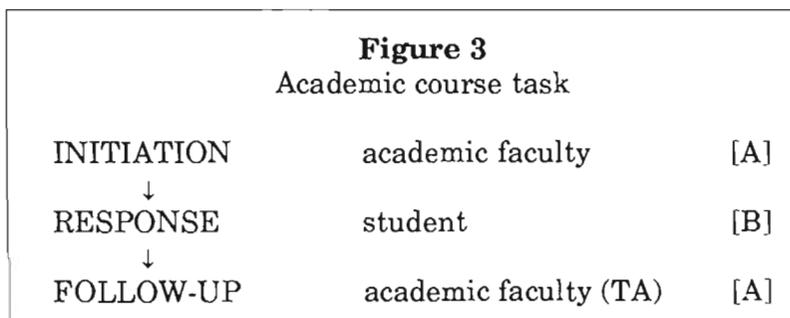
A writing test is a discourse exchange, although a discontinuous one, and this particular realization of the discourse exchange has the expected sequence:



The writing assessment which uses a discipline-related task is a three-way exchange:



This participant structure contrasts with that for essays written for academic courses, where exchange structure is typically two-way, with the professor setting and responding to the assignment, or perhaps supervising teaching assistants' responses:



When I was at the University of Edinburgh, I carried out a small scale study of the design parameters which faculty have in mind when they prepare written tests in their own discipline, and of the criteria they use when scoring the writing produced by the graduate students in their courses. Only one of the 24 faculty members I asked said writing was never used in the course. Seventy percent of the faculty used writing for most or all of their examination requirements at the graduate level.

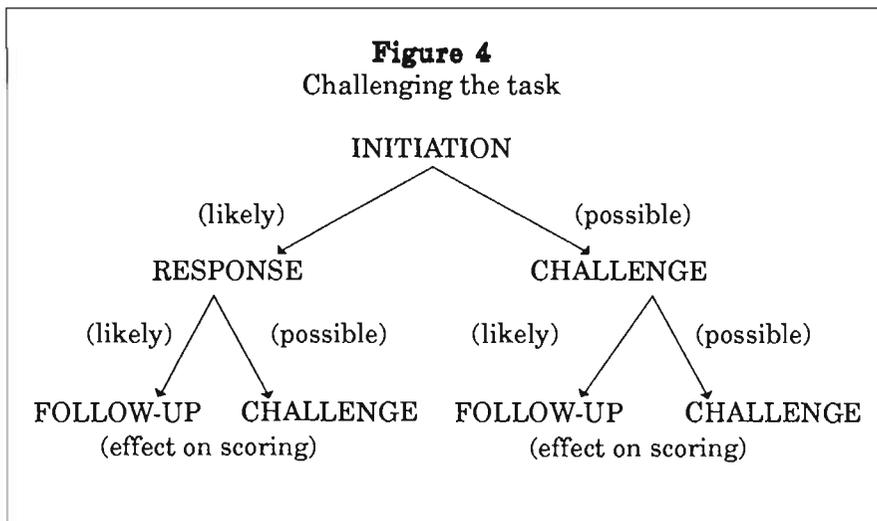
Faculty cited "avoidance of ambiguity" as the most common design feature they used for their essay examination questions, and the next most frequent design factor had to do with finding out what students know. Here faculty talked about **allowing** students to **show the ability to apply knowledge to problems in the discipline**; they talked about getting students to **show discrimination**; they talked about designing tasks that required students to **discuss or argue**, not just display knowledge.

Certainly, clarity was central, as was the ability to take information, think about it, integrate it, and reshape it into knowledge. But then I turned back to my data, to the international writing assessment context, and studied the essays students had written. I began to believe I was seeing some qualitative explanations for the statistically ambiguous answer to my question of whether questions located in a student's areas of knowledge allow her to look like a better writer than general questions.

The first and more obvious of these explanations is that any test with only six discipline-related choices does not offer enough options to give every candidate a task in an appropriate area. For example, in the technology area the task required the student to write about the properties of metals appropriate for one specific engineering use. Clearly, there are technical fields where individuals do not work with metals in the raw state. In attempting to design tasks which would "work" for as many candidates as possible, the test designers had been forced to choose topics mastered rather early in the learners' development of disciplinary knowledge. Some older candidates had forgotten the broader basic content of their freshman and sophomore years. But as I looked at essays and found textual evidence for lack of content knowledge or loss of content knowledge, I saw something else that began to fascinate me. Some of the discourse exchanges were "challenges," a

term first used by Labov and Fanshel (1977) in describing a discourse function in sociological, psychological terms. Instead of supplying the predicted response, some test takers challenged the question itself.

The addition of the challenge to the discourse between faculty member and student significantly expands its potential:



In a discourse exchange on a writing assessment, the respondent to a question, as a participant in a discourse, has a choice. She may give the predicted, what I call the “unmarked,” response, or she may “challenge”—take issue with the validity of the question. The challenge is a “marked” response, what in some fields might be called an “abnormal” response, and here as in many other fields of the human sciences we can learn a good deal about normal behavior or performance from the abnormal.

Because many writers, sensible of the pragmatic function of the genre of the writing test before them, edit themselves out of their texts—or, and perhaps more commonly, with new arrivals into our academic community, edit a new self into the text, much of the interaction of the individual writer and the task that interests me

is not amenable to close investigation. Bartholomae (1985) looked at how new undergraduate writers use the written language to locate themselves within an academic discourse. Taking the position that to write at the university level involves establishing authority for oneself, Bartholomae saw some students taking an “outside” position—creating a “naive” discourse—where they establish only the authority to be present and the discourse says simply: **I am here. This is who I am.** At a transitional level, students mimic the sound and texture of academic prose without there being any recognizable interpretive academic activity under way. This is the attempt to claim the authority of **membership**. Many writers taking our ESL writing assessments will fall into this middle category, will be at the stage of taking on the protective coloration of their discipline, and thus it will be difficult for us to see their personal imprint upon the texts they create. But, says Bartholomae, more advanced students are those who “establish their authority as **writers**, who claim authority by placing themselves at once within and against a discourse community, working consciously to claim a territory of knowledge as their own”(158). Looking at the writing that such students do on writing assessments such as the ELTS, where disciplinary knowledge is deliberately introduced as a factor by the testing agency, may help us understand both test and writer better.

When we look at how nonnative graduate students write about discipline-related content, we are at the intersection of two kinds of knowledge: knowing how to write, and knowing a discipline. Many of the texts I studied appeared to be by students who were “transitional” both in language and content knowledge. Constrained by the awesome implications of the test they were taking, it is unsurprising if they generally focussed on “psyching out” the test, the task, and the reader’s expectations. They wrote to camouflage, not to communicate.

Attempting to write with the voice of disciplinary membership, handicapped too by writing in a language other than their own, they did not offer the reader their voice, their sense of self, their authority. These, then, constitute what I am calling here my “unmarked” responses. Less competent writers, however, who had not yet learned how to distance themselves through rhetorical,

stylistic, and lexical shaping, were able only to establish the authority of **presence** (in Bartholomae's terms). When the interweaving of who they were and what they knew with the demands of the task became impossible for them, they were able only to protest or apologize.

1. I have no idea about "Green Revolution" but I think it is how to increase food supply.
2. My field of work is a computer science, so I'm afraid I can not exactly and rightly discuss the advantages and disadvantages of nuclear energy.
3. I have to recognize I don't exactly know how a factory is organized and then I cannot give a well opinion.

These are "marked" responses in my terms, but their lack of competence does not permit them the flavor of challenge. It is the other kind of "marked" responses, the papers where the student, more or less skilfully, but in my reading always authoritatively, takes a critical rather than an imitative stance in taking and shaping the task to herself, that I refer to as **challenging the task**. Challenging the task is not a new concept: John Swales (personal communication) recalls his days at Cambridge in the late 1950s, where the tradition even at the undergraduate level was to push against a language that would render one like everyone else, and where whenever the student was aware of a possible deficiency in content knowledge, the traditional strategy was "Don't **answer** the question, **question** the question." But I suggest that by focussing in on all challenges we can learn more about tasks as they affect the wider population of student writers.

In oral discourse a challenge is an interruption in the normal flow of the social conventions of conversation, and similarly in written discourse a challenge is an interruption in the normal flow the social conventions of textualization, in the making of meaning between writer and reader. Challenges do not conform to the discourse expectations, and as Figure 4 indicates, while unmarked responses are unlikely to lead to a challenge, a marked response is considerably more likely to do so. In challenging the task, then, in pushing against the conventions, especially in a situation where as

much depends on the outcome as this one, the writer is taking a risk. As Swales (in press) has shown, writers at the center of their disciplines have a firm understanding of the strength and reasonableness of the claims they make, that is, they have self-confidence in their powers of judgment, in their ability to see a problem and say what needs to be done about it. But they have, in addition, a strong sense of their location in that discipline, of how far they can go in transcending the "rules" of their community. To "work self-consciously against the common code," as Bartholomae has referred to it, is a difficult and dangerous activity. Not only is it subversive, it can be self-destructive.

Myers's (1985) study of two biologists writing grant proposals for work in somewhat new fields shows how they needed to establish a claim to authority in that field, transferring only some of their authority from their usual field. The rhetorical demands on them as they interpreted their work for a disciplinary community that had not yet awarded them membership were far greater than usual. Myers describes the tentativeness with which they moved beyond the realm of their established authority, hedging their claims and supporting their right to make them with "face establishing" moves (Brown & Levinson, 1977). Taking this risk, one was successful and one was unsuccessful. When a student, even a graduate student, takes this kind of risk on a test which will affect her future, we must think about whether she is displaying authority or masking incompetence. When a graduate student who is a nonnative speaker does so, we must ask what has compelled her to push against the test.

I will look at just two examples of challenges in my test data. Both examples come from the set of responses to a prompt in the field of physics (given in the Appendix). As a close look at that prompt will show, writers are instructed to take a specific **perspective** (Hamp-Lyons, 1987), that is, to argue from a pre-specified position, and further, the position specified for them is one contrary to that actually held by many people. The data showed that more writers challenged on this prompt than on any other. I find in this point a strong indication of the interweaving of influences of writer and prompt, as I later find evidence for the still more complex interweaving of writer, prompt, and reader.

In my first example (Appendix, A), the writer establishes authority with **"I am a scientist"** but at the same time resists the implication that is set up by the question, that *as a scientist* he will inevitably take a perspective which supports the position stated. This is a very difficult rhetorical move, and he manages it successfully, it seems to me, with the strong opening phrase **"Even though"** used as an extratextual reference to the prompt. The writer picks up on the part of the prompt which was intended to provide the position for him to counter, and instead of countering he argues in its favor. From the first clause on, however, he makes no further reference to his authority as a scientist, using arguments and style that would just as likely come from any other writer. Having established "face" as a scientist, the writer does not insist on that role in the rest of his essay. Further, from that point on his answer is not marked as a challenge *except* that the position he argues opposes that set up for him.

In the second example (Appendix, B), the writer makes no apology. Her answer is an argument. She establishes her definition of what it means to be a scientist, then uses her definition to support her position that the question is not one to be asked of the scientist. When she says **"For me, the question above does not make sense,"** she speaks from the role she is adopting for the task.

Throughout her essay she maintains the role of scientist, as we can see by following the chain of lexis relating to the objective nature of scientific inquiry ("can be measured"; "know"; "facts"; "proper information") and she maintains her sense of distance from the issue ("**only** can develop"; "**only** can give"). We can note a stress on knowing and knowledge, juxtaposed with awareness of the limits of knowledge and the need for a human responsibility beyond absolute facts. These are coupled with a more subtle expression of the complexity of the issues: first she uses two contrasting adjectives to suggest the balancing of beneficial and harmful effects, and second, two contrasting passive verb forms to suggest the contradiction of enjoying and being harmed at the same time. She is claiming her authority, her privilege to speak, by placing herself, as Bartholomae would say, "within and against" a discourse (1985, p. 158), and in so doing she is appropriating the task, remaking it as her own. Through much of the text her voice is confident, her

indignation is righteous. The move from “Therefore” in paragraph two to “To me” in paragraph three suggests, however, an awareness that what was “proved” in the foregoing was not really proved, but left something to be desired. From my experience of students from a range of cultures and differing exposures to the academic conventions implicit in the test, I would suggest that her unease betrays a sense that her forthright approach to the topic is not quite what the decision makers are looking for here, that perhaps something more, or different, is expected of her, although she doesn’t know what.

In the examples, the first writer takes a small risk in taking up an opposing position, but perhaps he is more aware of the risk he takes than we would think. Perhaps he is aware both of the task and its unreasonableness, and of the judges and their powers and predilections. If he sees them as liberal intellectuals, he is not far off the mark. If he thinks that the unreasonableness of the task and the probability that the essay scorers will be English teachers are related, he is even closer. The second writer, on the other hand, displays no rhetorical sleight-of-hand. She deals with the issue from within the discipline, and expects the reader to meet her within the discipline. If we do not do so, can we say it is she, and not we, who is at fault?

I would like to argue that neither of these writers displays incompetence as a writer, that both know who they are and what they are doing. The incompetent writing that we looked at earlier was marked particularly by the appeal to the reader, by an inability to stay within the “rules” of this game, which require that the writer writes as though unaware that this text is to be evaluatively, not collegially, read. Neither of these more able writers does this. But the first writer is aware of his reader, of his rhetorical context. He says in effect: “I am a scientist—but think of me as one of you.” His is an appeal to the emotions. The second writer says: “The initiation of this discourse is grounded in invalid assumptions, and I must draw your attention to that fact.” It is an appeal to reason. In this she is apparently unaware of her reader, her rhetorical context.

Here I must briefly call attention to the three-way nature of this discourse exchange. It is not only writer and prompt that interact to create the “response,” marked or marked, on an essay test: the

reader, that human judge, is truly a discourse participant. I have argued elsewhere (Cooper & Hamp-Lyons, 1989) that in large measure readers create the discourse anew as they read. If readers experience the discourse as discontinuous, as disrupted, we may predict that their responses will be negative. This possibility is represented in Fig. 4 by the *challenge* arrow after the writer's *challenge*. I suggest that the reader's inclination to challenge the writer's challenge may take the form of a low score, if the challenge is one the reader does not value. In contrast, as is implicit in the behavior of those Cambridge undergraduates of the 1950s, the "smart" challenge, the ability to take on the professor on his own terms and hold one's own, is often highly-valued and well rewarded.

What are the characteristics of the first example that led essay evaluators, English teachers all, to assign scores of 8 and 9 on a 9-point scale? My discussions with English specialists have shown, over and over, that comments focus on its language fluency. The writer is persuasive and controls persuasive language. And what are the failings of the second sample that led those same evaluators to assign scores of 4 and 5 on the same scale? Again, my discussions with English specialists have shown repeatedly that comments focus on incorrect lexis and poor spelling. The piece is not persuasive, does not engage the reader—and in not engaging the reader it leaves her free to notice the language weaknesses.

What makes the first piece fully competent or very close to it, and the second piece halfway to competence at best, for these judges? I would suggest that the emotional quality of the challenge in the first plays a large part in the readers' judgments, and that the rational quality of the challenge in the second similarly plays a large part in readers' judgments. I would further suggest that the first essay is not twice as good as the second, and that this is a case where readers' judgments are little based in the actual strength of the ideas or arguments the writers are making. Several professors in disciplines outside English/ESL have looked at these two pieces together, and have found them roughly comparable in writing quality, where the definition of "quality" was left open for them to make for themselves.

## Conclusion

I believe that the focus on challenges, as exemplified by the two cases discussed above, has opened for me a whole range of insights and issues which can inform the study of prompts more generally. It has reiterated for me the critical significance of the selection and wording of prompts, not only in general but also in relation to the needs and background knowledge of the anticipated test audience. It has also led me back to basic issues such as the impact of readers on texts, and judgments of text quality. Once again I realize how improbable it is that we will ever pin down prompt difficulty, since so much of the difficulty is not in the prompt or the writer, but in the reader.

Without accompanying studies of students as they encounter written examinations we shall remain unsure of much that is important to an understanding of essay tests: how (indeed if) students read the prompt, how they establish salience for elements of the prompt, what makes them choose one prompt over another where a choice is given, how they decide what persona to present to the reader—all of this lies behind our numerical data. Since the many studies of topic effect have not led to consensus in the writing assessment community, a shift of focus seems called for. The study of challenges suggests that for a genuine understanding of what writers do on essay tests, and why, we will have to go much further. One possibility is the case study: an examination of the candidate's writing development prior to the test. Such work is long overdue.

## References

- Bartholomae, D. (1985). Inventing the university. In M. Rose (Ed.), *When a writer can't write: Writer's block and other composing process problems* (pp. 134-165). NY: Guilford.
- Bridgeman, B., & Carlson, S. (1983). *Survey of academic writing tasks required of graduate and undergraduate foreign students*. Princeton, NJ: Educational Testing Service.
- Brown, P., & Levinson, S.C. (1987). *Politeness* (2nd ed.). Cambridge: Cambridge University Press.
- Carlson, S., Bridgeman, B., Camp, R., & Waanders, J. (1985). *Relationship of admission test scores to writing performance of native and nonnative speakers of English*. (TOEFL Research Report 19). Princeton, NJ: Educational Testing Service.

- Cooper, G., & Hamp-Lyons, L. (1989). *Looking in on essay readers*. Ann Arbor, Michigan: English Composition Board.
- Hamp-Lyons, L. (In press). Second language writing: Assessment issues. In B. Kroll (Ed.), *Second language writing: Issues and options*. New York: Cambridge University Press.
- Hamp-Lyons, L. (1986). Testing second language writing in academic settings. Unpublished doctoral dissertation, University of Edinburgh, Edinburgh, Scotland.
- Hamp-Lyons, L. (1987). Performance profiles for academic writing. In K. Bailey, R. Clifford, & T. Dale (Eds.), *Language testing research* (pp. 78-92). Monterey, CA: Defense Language Institute.
- Hamp-Lyons, L., & Prochnow, S. (1990). *Prompt difficulty, task type, and performance*. Paper presented at the Language Testing Research Colloquium, San Francisco, CA.
- Horowitz, D. (In press). ESL writing assignments: Contradictions and resolutions. In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts*. New York: Ablex.
- Horowitz, D. (1986). What professors actually require: Academic tasks for the ESL classroom. *TESOL Quarterly*, 20, 445-462.
- Labov, W., & Fanshel, D. (1977). *Therapeutic discussion: Psychotherapy as conversation*. New York: Academic Press.
- Myers, G. (1985). The social construction of two biologists' proposals. *Written Communication*, 2, 219-245.
- Reid, J.M. (1989). Quantitative differences in English prose written by Arabic, Chinese, Spanish, and English students. *Dissertation Abstracts International*, 50, 672-A. (University Microfilms No. 89-11, 827).
- Spaan, M. (1989, March). *Essay tests: What's in a prompt?* Paper presented at the TESOL Convention, San Antonio, TX.
- Swales, J. (In press). *English in academic and research settings: Genre analysis and its applications*. New York: Cambridge University Press.

## Appendix

### Prompt Question and Two Challenge Responses

*1. New scientific processes often meet with opposition because of the pollution they cause to the environment, an example of which ...  
As a scientist, how would you defend the continued use of such potentially harmful processes?*

#### *Example A*

Even though I am a scientist, I strongly consider the opposition to new scientific processes as a healthy action. I think that man's new experiments should compromise their experiences and the environment, in order to avoid damages to the present and future nature. Basically, no man or nation has the right to, in the name of a scientific progress, destroy their own habitat—the only one we have now.

Fortunately, though, I believe that man can find better ways to guarantee the environment and mankind preserved. Indeed, I also believe that new scientific processes can be done just to improve the quality of life on Earth, in despite of all economic interests involving this man's action all around the world. Even though someone can find this thesis completely utopic, I really trust it.

#### *Example B*

The physical facts of pollution can be measured by using scientific equipments, and scientists know the process of the facts. Engineers who know the scientific knowledge only can develop facilities which reduces this harmful processes. On the other hand, politicians and executives of companies have a force to decide the use of the beneficial but harmful processes. The decision must be or reflect the will of people who are enjoying and are harmed by the process.

Therefore, the scientists only can give people proper information about the process, and the engineers only can give people proper information about the technology and the cost of preventing the harmful effects.

For me, the question above does not make sence. The choice of continual use of such potentially harmful process or cutting off the use of the process does not depend on the scientist. Scientists want to know everything in a rational way. The knowledge obtained by this way is so repeatable and testable, or reliable, that this knowledge have a power. The way how we use the power is not on the responsibility of the scientists.

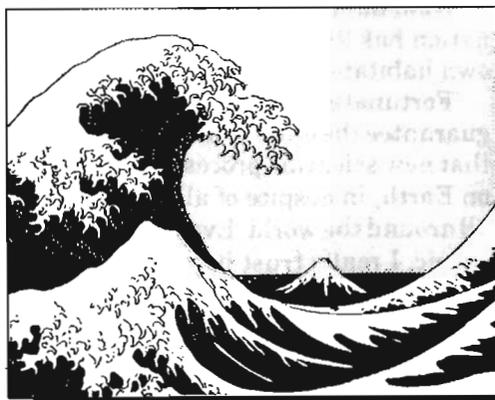
# CROSS CURRENTS

**AN INTERNATIONAL JOURNAL OF  
LANGUAGE TEACHING AND CROSS-CULTURAL COMMUNICATION**

*Cross Currents* is a semi-annual publication of the Language Institute of Japan (LIOJ) which provides a forum for the interdisciplinary exchange of ideas within the areas of cross-cultural communication and language skill acquisition and instruction. At *Cross Currents*, we are particularly committed to issues concerning both the theoretical and practical aspects of ESL/EFL instruction, cross-cultural training and learning, English language teaching as it applies to Japan, and English as an International Language.

**Recent contributors to  
*Cross Currents* include:**

Robert O'Neill  
Braj Kachru  
Alan Maley  
Walter Hirtle  
Richard Via  
Larry Smith  
Paul LaForge



SUBSCRIPTIONS	Single issue	1 year	2 years
<b>Inside Japan*</b>	¥1,300	¥2,600	¥5,150
<b>Outside Japan</b> — Individual Subscription	\$8.00	\$15.00	\$25.00
<b>Outside Japan</b> — Institutional Rate		\$19.50	\$35.00

\* Special rates available for JALT members. (See postal form in the Language Teacher.)

• **Japan Subscriptions:**  
School Book Service  
FI Building 4F  
1-26-5 Takadanobaba  
Shinjuku-ku, Tokyo 160  
03-200-4531

(By postal transfer to  
Tokyo postal account No.  
9-86192, or cash delivery -  
Genkin Kakitome.)

• Current Issues are also  
available in major Japanese  
bookstores through YOHAN  
in Japan.

• Back Issues are available  
from *Cross Currents*.

• **Outside Japan:**  
*Cross Currents*  
Language Institute  
of Japan  
4-14-1 Shiroyama  
Odawara, Kanagawa 250  
Japan

(Make checks payable  
to *Cross Currents* (LIOJ)  
drawn on a U.S. bank.)

## **Reading in a Foreign Language: Research and Pedagogy**

**Patricia L. Carrell**

*Graduate School, University of Akron*

This article, which originated as the author's plenary address to the 1989 JALT Conference in Okayama, Japan, surveys recent research in second or foreign language reading, especially that research conducted within the context of schema theory, and discusses the implications and applications of that research to classroom reading pedagogy. After an introduction to schema theory and its conceptualization of the reading process as an interaction of text-based and reader-based processes, the author discusses schema-theoretic reading research first from the perspective of content schemata, and then formal schemata, and then the interaction of the two types of schemata. Within each type of schemata, effects studies (showing the effects on second or foreign language reading comprehension of content and formal schemata) as well as training studies (showing the facilitation of second or foreign language reading by training the appropriate content and formal schemata) are reviewed. The author covers her own research, as well as that of others.

The second part of the article discusses the teaching implications and applications of the research to second or foreign language reading classrooms. Prereading activities, semantic mapping, "name-brand" reading methods, dialogue journals, the reading laboratory approach, content-centered approaches, text-mapping, and rhetorical approaches to reading are discussed.

### 外国語教育における読解：理論と実際

本論文は、1989年に岡山で開かれた全国語学教育学会の全国大会での講演原稿に基づくものであり、最近の第二言語もしくは外国語教育としての読解指導のうち、特に、スキーマ理論 (schema theory) の研究を中心に、その実践的な含みと教育現場への応用につき概観し、論ずるものである。本理論 (schema theory) につき概説し、読解過程における読み手の言語的、経験的な知識の相互作用に関しての概念的な側面を紹介した後、まず、テキストの提供する話題の種類と、その構文の種類に関してそれぞれ取り上げ、さらに両者の読解への相互作用に関する研究について論及する。また、(テキストの内容・話題とその構成・構文につ

いての知識が、第二言語もしくは外国語教育としての読解能力をいかに左右し得るものなのかといったことについての)研究及び、(両領域の適切な指導により、いかに学生が上達し得るかといったことを示す)訓練に関する研究についても、それぞれ、著者のみならず、他者の研究も取り上げ、概観する。

本論文の後半では、以上の概念が、第二言語もしくは外国語教育としての読解の授業にどのような示唆を与えるものであり、いかの実践的に応用し、活用し得るものであるか、といったことについて検討する。読解前の学習活動のいろいろ、意味連想作図法 (semantic mapping)、有名な読解指導諸方式 (“name-brand” reading methods)、教師—生徒問答交換法 (dialog journals)、図書カード読解方式 (the reading laboratory approach)、特定話題集中方式 (content-centered approaches)、要点抽出作図法 (text-mapping)、文章構成 (分析) 方式 (rhetorical approaches) についても検討を加える。

## 1. Introduction

Developments in *schema theory* since about 1977 (Adams & Collins, 1979; Anderson, 1978; Rumelhart, 1980; Rumelhart & Ortony, 1977) have had a pervasive influence on current thinking about text comprehension. Through an emphasis on the role that preexisting knowledge structures play in the mental processing of text, schema-theoretic approaches have revealed the complexities of the interactive, constructive processes necessary to comprehension (e.g., Schank & Abelson, 1977). Within a schema-theoretic framework, text comprehension—or, more specifically for the purposes of this paper, reading comprehension—is characterized as an interaction of text-based processes and knowledge-based processes, both related to the reader’s existing background knowledge or schemata (Adams & Collins, 1979; Anderson, 1977; Bartlett, 1932; Rumelhart, 1977, 1980; Rumelhart & Ortony, 1977). The idea has been put succinctly by Anderson, Reynolds, Schallert, and Goetz (1977, p. 369): “Every act of comprehension involves one’s knowledge of the world as well.”

In a 1983 article, Eisterhold and Carrell described how schema theory conceptualizes the interaction of text-based and knowledge-based processes, or as they are more commonly referred to, *bottom-*

*up* and *top-down* processing modes. This interaction of text and prior knowledge may be illustrated with the following mini-text:

Jane was invited to Jack's birthday party. She wondered if he would like a kite. She went to her room and shook her piggy-bank. It made no sound. (Charniak, 1972)

Take a moment and think about this text and the interpretation you have arrived at for it, and consider how you arrived at it. Because we do it so automatically and subconsciously, it is often difficult to make the process conscious and overt.

Upon encountering the first sentence, the incoming data presumably trigger a "birthday party" schema. The rest of the text is interpreted against that schema, which helps to flesh out the information left implicit in the text. We know that one part of the birthday party schema includes birthday presents; therefore, we assume that Jane's wondering whether Jack would like a kite is because she is trying to think of a suitable birthday present to give him, and that a kite would be a possible present. Further, from our subschema for gifts, we know that gifts are frequently purchased. Thus, we can make sense of the third sentence. Jane wants to buy Jack a birthday gift, and we know that a purchase requires money. Our schema for piggy banks tells us that this is often a place a young child saves money. (Other things about the text also suggest that Jane and Jack are young children—the names and the short sentences are typical of children's stories.) The fact that the piggy bank makes no sound, related to our knowledge that piggy banks generally contain coins and not paper money, tells us that there is probably no money in the piggy bank. Therefore, we conclude that Jane will have to solve her problem some other way. (Notice, incidentally, how many of these schemata and subschemata are potentially culturally-based, and cannot be assumed to be universal. Birthday parties may be widespread in many cultures, but we cannot conclude that every culture celebrates birthdays, or celebrates them with parties; the giving of birthday gifts may also be widespread, but need not be universal; and the buying of gifts rather than making them, or the giving of things one already possesses may also be culture-specific, or even specific to certain sub-cultures.)

As illustrated by discussion of this mini-text, schema-theory

research has shown that the most efficient processing of text is interactive—a combination of top-down and bottom-up processing modes (Rumelhart, 1977, 1980). Top-down processing is making predictions about the text based on prior experience or background knowledge, and then checking the text for confirmation or refutation of those predictions. Bottom-up processing is decoding individual linguistic units (e.g., phonemes, graphemes, words, on up to phrases and clauses) and then referring these analyzed units to one's background knowledge for confirmation of fit. Preexisting background knowledge and current predictions based on this knowledge are modified on the basis of information encountered in the text. Skilled readers constantly shift their mode of processing, accommodating to the demands of a particular text and a particular reading situation; less-skilled readers may tend to overrely on processes in one direction or the other, often producing negative effects on comprehension (Spiro, 1978, 1979).

In thinking about the interactive nature of reading between top-down and bottom-up processing modes, knowledge-based and text-based processes, it is important to bear in mind that this interaction involves all kinds of knowledge which the reader brings to the reading task. In addition to linguistic knowledge (i.e., linguistic schemata, including lexical, syntactic, semantic, as well as pragmatic knowledge), the reader also brings knowledge and beliefs about the world (i.e., content schemata) and knowledge about texts of different types and how they are typically organized or structured (i.e., formal schemata). We might also add knowledge about or beliefs about the reading process itself and personal goals in reading. The interactive nature of text processing, involving both top-down and bottom-up processes, occurs both within and across various levels of processing, from the lowest levels of feature, letter, and word recognition, to syntactic and propositional levels, to the highest, most global aspects of text and context. The interaction is not only between and across levels of processing within the reader, but also between the reader and the text—between levels of processing within the reader and the properties of the text at various levels of analysis.

In my own research, I have often found it useful to distinguish content schemata from formal schemata and to investigate the separate, distinct roles both of these kinds of knowledge play in

second language reading. I would like to make that distinction at this point, and to first discuss studies which have focused on content schemata—that is, studies which have demonstrated the effects of content schemata on second language reading comprehension, and training studies which have shown the efficacy of teaching relevant background content. Thereafter, I will turn to studies which have focused on formal schemata.

## 2. Content Schemata Studies

A number of second language, ESL studies have shown that prior background knowledge of the content domain of a text significantly affects reading comprehension of that text. This result has been demonstrated in particular for prior background knowledge of culture-specific text content.

The seminal study of this type was done by Steffensen, Joag-dev, and Anderson in 1979. In that study, two groups of subjects with different cultural heritages were investigated—a group of Asian Indians living in the United States and group of Americans. Each subject was asked to read and recall two personal letters, both of which were constructed with similar rhetorical organization. However, the cultural content of the two letters differed: one described a traditional Indian wedding, the other a traditional American wedding. It was assumed that all adult members of a society would have a well-developed system of background knowledge about the marriage customs of their own culture and a relative lack of knowledge about the marriage customs of more distant cultures. This is exactly what Steffensen et al. found. Although the Indian group overall read the texts more slowly and recalled less than the American group—a reflection of the fact that they were performing in English, their second language, rather than their native language—each group read the material dealing with their own cultural background faster and recalled more of the content than they did of the text from the more distant culture. Furthermore, members of the culture provided appropriate cultural elaborations; non-members provided inappropriate cultural distortions—frequently outright intrusions from their own culture. In short, the study showed a clear and profound influence of cultural content schemata on reading comprehension.

Johnson (1981) also investigated content schemata with two authentic folktales and two groups of readers—a group of Iranian students studying in the United States and a group of Americans. Both groups read a Mullah Nasr-el-Din story from Iranian folklore and a story about Buffalo Bill from American folklore. Both folktales “contained similar motifs which were culturally distinct yet were equivalent in plot construction” (p. 170). Johnson’s results were much like those of Steffensen et al.—that is, superior performance on a given text by members of the cultural group, poorer performance by non-members of the cultural group—thus clearly showing strong effects of cultural content schemata.

An interesting side aspect to the Johnson study was that she also manipulated the syntactic and semantic complexity of each text. Half of the subjects in each group read the stories in adapted or simplified English, the other half read unadapted versions of the same stories. Her results showed that the cultural origin of the story had greater effect on the comprehension of the ESL students than did the level of syntactic and semantic complexity of the text. That is, the Iranians performed better on a text from their native culture than on a text from American culture, and there were no differences in their performance on the Iranian text between the simplified and unsimplified versions. The Iranians did perform better on the simplified American text than on the unsimplified version. Johnson concluded that if ESL readers have the appropriate content schema for a text they can cope equally well with unadapted, syntactically and semantically unsimplified texts. Similar studies of content schemata have been conducted by Cabello (1984) and Haus and Levine (1985).

Some of the work on the effects of content schemata on reading comprehension has been conducted in the area of English for Specific Purposes (e.g., English for science, technology, business, or English for academic purposes). The general ideas behind this research are: first, as Widdowson (1979) has observed, different disciplines (such as physics) constitute subcultures of their own into which readers are enculturated; second, texts and modes of communicating via texts in each discipline or subculture may vary; and third, material from a familiar discipline or subculture is easier to read and understand than linguistically comparable material from a less familiar discipline (Alderson & Urquhart,

1985; Cohen, et al., 1976; Mohammed & Swales, 1984; Strother & Ulijin, 1987).

In fact, Alderson and Urquhart's work with reading English for Academic Purposes has led them to question the traditional notion behind the selection of texts for testing purposes—namely the aim of selecting texts which are sufficiently “general” in order to avoid favoring any particular group of students. Underlying this traditional position is obviously a belief that certain texts *will* favor particular groups, presumably because of the background knowledge available to them. However, Alderson and Urquhart point out that such general texts may not be appropriate measures of EFL reading comprehension. In an empirical study of English for Academic Purposes, with students from different disciplines reading discipline-specific texts as well as so-called “general” texts, they found (a) that students from a particular discipline performed better on tests based on texts taken from their own subject discipline than did students from other disciplines (that is, students appear to be advantaged by taking a test on a text in a familiar content area); (b) that students from certain disciplines found the so-called “general” texts easier than did students from other disciplines (that is, the texts were not “general” across all discipline groups, and, in fact, Alderson & Urquhart end up questioning the existence of truly “general” texts which would have to be so neutral in content and cultural/disciplinary assumptions that they would not, in some way, favor a particular group; and (c) that these “general” texts underestimated the reading ability of science and engineering students when compared to their reading ability on texts in their disciplines.

Alderson and Urquhart concluded that it is the more specialized, not the more generalized texts, which may be the best tests of a reader's EFL reading ability. For second or foreign language readers, many of whom have much more limited skills for extracting information from texts and whose second or foreign language reading skills have been developed in specific discipline contexts, they argue that inability to perform successfully on so-called “general” texts may not be indicative of their abilities on texts in their own specialities.

I am aware of only one second language, cross-cultural reading study which has investigated and shown the facilitative effects of

actually training or teaching relevant content schemata (Floyd & Carrell, 1987). This study raised the pedagogical question "Can we improve ESL students' reading comprehension by helping them to build background knowledge on the topic prior to reading?" Our results suggested an affirmative answer. We used pre- and post-tests with experimental and control groups of intermediate-level ESL students (TOEFL 400-499), half of each group receiving syntactically more complex versions of the test passages than the other half. The experimental group was taught appropriate cultural background information between the pre- and post-testing. Results showed that by providing the experimental groups of students with relevant first-hand experiential knowledge, reading comprehension—as measured both by objective multiple choice questions and by a free written recall—was facilitated. Similar to Johnson's earlier results with more advanced ESL subjects, our results with these intermediate-level ESL subjects showed that cultural background knowledge was more of a determining component of reading comprehension than was syntactic complexity. The level of syntactic complexity, in fact, had no significant effect on either way of measuring reading comprehension.

### 3. Formal Schemata Studies

A number of research studies have provided empirical evidence that the rhetorical organization of a text interacts with the reader's formal schema—that is, the reader's background knowledge of, and experience with, textual organization—to affect reading comprehension. This effect of text structure on reading comprehension has been shown for both narrative and expository texts. In a 1984 study (Carrell, 1984a), I investigated the effects of a simple narrative formal schema on reading in ESL. Earlier findings of Jean Mandler (1978; Johnson & Mandler, 1980; Mandler & Johnson, 1977) had demonstrated that native English-speaking children as early as the first grade have acquired a story schema and use it to organize their comprehension and recall of simple narrative stories. In fact, in her cross-cultural research with the Vai in Liberia, Mandler had suggested that such a narrative schema may be universal. In my study I found differences in the quantity and temporal sequence of ESL readers' recalls between those reading

standard versions of simple stories and those reading interleaved versions of the same stories. Quantity of recall was enhanced when the story's rhetorical organization conformed to the simple story schema—one well-structured episode followed by another. When stories violated the story schema, the quantity of recall was reduced, and the temporal sequencing of the readers' recalls tended to reflect the story schematic order rather than the temporal order of presentation in the story. These differences reflect the effects of the simple story schema on the comprehension of simple English stories by ESL readers.

In the realm of expository prose, Hinds (1983a, 1983b) has compared Japanese and English speakers reading, in their respective native languages, texts with a typical Japanese rhetorical structure. His findings show that not only is the Japanese structure generally more difficult for the English readers, but that particular aspects of that rhetorical organization are extraordinarily problematic for them, especially in delayed recall. He concludes that the traditional *ki-sho-ten ketsu* pattern of contemporary Japanese expository prose is more difficult for English readers because of its absence in English expository prose. That is, native English readers lack the appropriate formal schema against which to process the Japanese rhetorical pattern.

In another 1984 study (Carrell, 1984b), I found effects of four different English expository patterns on the reading recall of ESL readers of various language backgrounds. That study showed that the more tightly organized patterns of comparison, causation, and problem-solution generally facilitated the recall of specific ideas from a text to a greater extent than the more loosely organized collection of descriptions pattern. In this finding, ESL readers generally appeared to be similar to the native English readers tested by Bonnie Meyer and Roy Freedle (Meyer & Freedle, 1984). However, that study also found significant differences among the language groups tested (Arabic, Spanish, and "Oriental") as to which English expository patterns facilitated greater recall. For example, the Spanish group was most like the total group, finding the collection of descriptions (C of D) type of organization far less facilitative of recall than any of the other three types of organization, which were all comparable for them. The Arabic group found the causation (C/E) type the least facilitative, the compare and

contrast (C/C) type the most facilitative, while the "Oriental" group (predominantly Korean plus a few Chinese) found the causation and problem-solution (P/S) types the most facilitative, and the comparison and collection of descriptions about equally less facilitative. Urquhart (1984) and Benedetto (1984, 1985) have conducted similar studies that show the significant effects of expository text structure or organization on second language reading comprehension.

In a controlled training study (Carrell, 1985), I set out to answer the question, "Can we facilitate ESL reading by explicit teaching of text structure?" In that study we taught intermediate-level ESL students for only one week about top-level rhetorical organization using the types identified by Bonnie Meyer, C/C, C/E, P/S, C of D. The idea was to teach the students about these types of textual organization so that they could use this information as a strategy with which to organize their interactive reading of a text and, later, their recall of it. The training experiment yielded promising results, demonstrating that explicit, overt teaching about the top-level rhetorical organization of texts can facilitate ESL students' reading comprehension as measured by quantity of information recalled. Qualitative analyses showed that the teaching facilitated recall of supporting detail as well as of major topics and subtopics. In addition, the effects of the training were evident for as long as three weeks after the teaching had ended. Further, student reaction to the teaching was extremely positive. Students expressed the view that they had learned a helpful technique and they expressed more confidence in themselves as ESL readers. One very shy student said that most of his life he had hated reading because he never knew what he was looking for, but that now reading made sense to him.

#### **4. Content Schemata versus Formal Schemata**

In one final study (Carrell, 1987), I examined the simultaneous effects of both culture-specific content schemata and formal schemata. I wanted to see which was the more important in affecting reading comprehension. In that study, high-intermediate ESL students (TOEFL 450-525) read, recalled, and answered questions about each of two texts. For each of the two groups of readers—

students of Muslim and Catholic/Spanish backgrounds—one text had culturally familiar content, the other culturally unfamiliar content. The two texts used were fictionalized historical biographies, one about “Saint Catherine” and one about “Ali Affani.”

The hypothesis was that the Muslim students would be relatively familiar with the content of the Ali Affani passage, and relatively unfamiliar with the content of the Saint Catherine passage, and vice-versa, the Catholic/Spanish students would be relatively familiar with the content of the Saint Catherine passage, but relatively unfamiliar with the content of the Ali Affani passage. In addition, within each group, one half of the subjects read the texts in a familiar, well-organized rhetorical format; the other half read the texts in an unfamiliar, altered rhetorical format.

Results showed that the conditions expected to yield good reading comprehension did so—namely familiar content and familiar rhetorical form. Texts with familiar content and familiar rhetorical form were easy. Similarly, the conditions expected to yield poor reading comprehension also did so—namely unfamiliar content and unfamiliar rhetorical form. Texts with both unfamiliar content and an unfamiliar rhetorical form were difficult.

More interestingly, the results for the “mixed” conditions—that is, familiar content in unfamiliar rhetorical form and unfamiliar content in familiar rhetorical form—showed that content schemata affected reading comprehension to a greater extent than formal schemata. Texts with familiar content, even if in unfamiliar rhetorical form, are relatively easier than texts in familiar rhetorical form but with unfamiliar content. Thus, within the limitations of the particular study, the overall finding seemed to be that when both content and rhetorical form are factors in ESL reading comprehension, content is generally more important than form. When both form and content are familiar, the reading is easy; when both form and content are unfamiliar, the reading is difficult. When either form or content is unfamiliar, unfamiliar content poses more difficulties for the reader than unfamiliar form. However, rhetorical form was a significant factor, more important than content, in the comprehension of the top-level episodic structure of the text and in the comprehension of event sequences and temporal relationships among events. In other words, both content and form play significant but different roles in the comprehension of text.

## 5. Implications: The Teaching of Reading

What are the implications for ESL reading teachers of this research on content and formal schemata? I shall first consider the teaching implications for content schemata.

Without wanting to overstate the teaching implications based on only the one training study of cultural-content schemata (Floyd & Carrell, 1987), I do not believe it is an overstatement to say that this study, plus the effects studies which preceded it, suggests that in the EFL/ESL reading classroom cultural content is of the utmost importance, and that it often must be explicitly taught. Teachers of EFL/ESL reading need to be teachers of culturally appropriate content, facilitators of the acquisition of appropriate cultural content schemata. Kathleen Stevens's observation about first-language reading teachers applies equally, if not more so, to EFL/ESL reading teachers: "A teacher of reading might thus be viewed as a teacher of relevant information as well as a teacher of reading skills" (1982, p. 328). If students are to develop proficiency in reading culturally unfamiliar material, the EFL/ESL teacher must provide the student with the appropriate cultural schemata s/he is lacking, and, more importantly if this training is to generalize to other reading situations outside the ESL classroom, the teacher must also teach the student how to build bridges between existing knowledge and new knowledge.

Failing definitive pedagogical research on which teaching methods work best in building background knowledge—for example, direct versus symbolic experiences, direct versus incidental instruction, inductive versus deductive instruction—the best the classroom reading teacher can do is to experiment with a number of prereading activities. Direct teaching of appropriate background knowledge can be accomplished through lectures, or various other types of prereading activities, including: viewing movies, slides, pictures; field trips; demonstrations; real-life experiences; class discussions or debates; plays, skits, and other role-play activities; text previewing; introduction and discussion of the key concepts and key vocabulary to be encountered in the text and the schemata surrounding key concepts; and even prior reading of related texts. Until pedagogical research tells us otherwise, it is probably best to

use a wide variety of prereading activities in varying combinations.

Of particular relevance for EFL/ESL readers at lower levels of proficiency and with limited vocabularies—for whom meaning tends to break down at the word level—are prereading activities involving key-word or key-concept association tasks. Pearson and Johnson (1978) propose the use of word association tasks in instructional settings to yield a diagnosis of what students already know and what they need to know about a key concept. The teacher throws out to a class a key concept or key word and asks the students to volunteer their associations on the word. Initial associations made by students may be of different types (superordinates, subordinates, attributes, definitions, synonyms, antonyms, contradictories, contraries, reverses, personal experiences, or even similar-sounding words). As students volunteer these associations the teacher writes them on the blackboard or overhead projector. The teacher and the students may go even further and organize the associations into a semantic map (Johnson & Pearson, 1978).

Reflection on these associations may then form the basis for further class discussion. Langer (1981) has found that through such class discussion students may significantly enrich their networks of associations and their vocabularies. In attempting to get students to “stretch” their concepts, Pearson and Johnson (1978) and Pearson and Spiro (1982) encourage the teacher to use analogies, comparisons, even metaphors to build bridges between what the students already know about a concept and what they may need to know in order to read and understand a particular text. Obviously, it is also helpful for the teacher to offer several examples of the new concept, as well as several examples of what it is not, so students develop a sense of the boundaries of the concept.

Such semantic mapping can be used jointly as both a pre- and post-reading activity. Johnson, Pittelman, and Heimlich (1986), and Heimlich and Pittelman (1986) claim that using a semantic map as a post-reading, follow-up activity “affords students the opportunity to recall, organize, and graphically represent the pertinent information read” (1986, p. 781). It further provides students with an opportunity to compare information gained from reading to information they possessed prior to reading, and to integrate new information.

Several organized approaches and methods have been proposed

in the literature for facilitating reading through activation of background knowledge. By “organized” I mean methods that have been given a label or name, have been at least somewhat codified, and are already published and accessible in the pedagogical literature. I am referring to such approaches as: The Language Experience Approach (LEA; Hall, 1981; Rigg, 1981; Stauffer, 1980); Extending Concepts through Language Activities (ECOLA; Smith-Burke, 1980); Directed Reading-Thinking Activity (DRTA; Stauffer, 1980); Experience-Text-Relationship Method (ETR; Au, 1979); PreReading Plan (PReP; Langer, 1980, 1981); and, finally, Survey-Question-Read-Recite-Review Method (SQ3R; Robinson, 1941). Without going into the details of each of these methods, I shall merely summarize what they all seem to have in common. For further information and the details of each method, interested readers may consult the literature, or refer to Barnitz (1985), which has an extended discussion of each method.

All of these methods train the reader to *do* something *before* reading in order to activate appropriate background knowledge—either creating the text themselves (LEA), setting a communication purpose for reading (ECOLA), predicting what a text will be about (DRTA), sharing prior experiences on the topic (ETR), free associating on the topic (PReP), or surveying the text (SQ3R). This prior activation of background knowledge also gives the reader a purpose for reading.

In addition, all of these methods have the reader read the text against the background of the activated knowledge. And, finally, they all have the reader *do* something *after* reading to synthesize the new information gained from the text with their prior knowledge—for example, discussing the text (LEA, SQ3R), writing their interpretations (ECOLA, SQ3R), reviewing the text to confirm hypotheses or prove conclusions (DRTA), relating text content to prior knowledge (ETR), or reformulating knowledge (PReP). Thus, all of these methods foster the building and activating of appropriate prior background schemata, and the integration of new information with old information. One or more of these methods are applicable to every proficiency level.

Another method which has been used successfully with more advanced ESL students, those able to communicate in writing, is the use of dialog journals (Steffensen, 1988). Students are assigned

an extensive piece of reading—for example, a novel anticipated to be somewhat unfamiliar to students in terms of its cultural content schemata. Students are assigned to do a prescribed amount of reading per week, and engage in a dialog with the teacher about the reading via the dialog journal. The teacher responds in writing. The students are encouraged to identify problem areas in the reading and to express their confusions or lack of understanding in their journals. In her responses, the teacher gently leads the student to see the cultural assumptions underlying the text, and how these may differ from the students' own cultural assumptions.

One of the strengths of the journal approach is that it allows for individual differences among readers interacting with the same text. It would be particularly useful in a class of students with heterogenous first language backgrounds and individual sources of cultural interference with the text. This approach would also work well with a homogeneous first language group where the teacher can anticipate common sources of cultural interference, and use these as a basis for class or group work.

Another way of developing a program which allows for a high degree of individualization in terms of interests, skills, and potential rates of progress is the reading lab class (Stoller, 1986). In this structured class, students are required to do as much reading as possible, but they make their own choices of reading material from a wide selection of appropriate texts. A reading lab has the additional advantages of allowing each student to progress at his or her own rate, and developing content schemata in some area of interest. Its disadvantages are that it presupposes a substantial library of materials, at a variety of difficulty levels, and, since there are no texts that every student reads, it limits the kinds of group work that can be done, and tends to isolate reading from other parts of the curriculum—or any other shared experiences. For students who are not strong self-motivators and who have not developed the habit of reading in depth in their own languages, this may be a major problem.

A final way of organizing second language reading programs is the content-centered approach (Eskey & Grabe, 1988). Variations of this basic theme include English for specific purposes (ESP) courses for particular academic or occupational groups, so-called "adjunct" courses attached to other academic courses (e.g., courses

designed to aid non-native speakers enrolled in a psychology or an engineering course), and so-called “sheltered” courses, which are limited to, and tailored to the special needs of nonnative speakers in academic programs (e.g., an introduction to English literature only for international students in an ESL setting).

A feature common to these courses is that they attempt to provide what Grabe (1986) has called a “critical mass” of information on a subject for the class as a whole to explore in depth, and this in turn provides a natural occasion for reading extensively in that subject. Appropriate pre- and post-reading work also emerges naturally in the form of introductory lectures or films, ongoing discussions of the subject matter, and, following the reading, the production of oral or written presentations, which provide real motivation for reading about the subject. Student interest is stimulated by classroom give and take, and there is a natural mixing of skills—listening, speaking, reading, and writing—as the students pursue a common intellectual goal. Reading is no longer isolated; it simply becomes an integral part of the normal educational process. It is not taught as an end in itself, but as a means to learning more about a subject.

Turning now to schemata, we find that a number of instructional strategies involving text-mapping have evolved recently to help make readers aware of the rhetorical structure of texts—the way the content of a text is organized. These strategies are also intended to help readers use their knowledge about the rhetorical organization of a text to guide and organize their interaction with the text. Although these strategies have arisen from research on text analysis of both expository and narrative texts, the discussion here will be limited to expository, or informational, texts.

Generally speaking, text-mapping involves selecting key content from an expository passage and representing it in some sort of visual display (boxes, circles, connecting lines, tree diagrams, etc.) in which the relationships among the key ideas are made visually explicit. Four such ideas have all been successfully used as instructional tools: “networking” (Dansereau, et al., 1979), “mapping” (Anderson, 1978), “flowcharting” (Geva, 1980, 1983), and “top-level rhetorical structures” (Meyer, 1975; Bartlett, 1978; Carrell, 1985). Students use text cues to define the fundamental relationships as they manifest themselves in expository text. All of these techniques

use a wide variety of prereading activities in varying combinations.

Of particular relevance for EFL/ESL readers at lower levels of proficiency and with limited vocabularies—for whom meaning tends to break down at the word level—are prereading activities involving key-word or key-concept association tasks. Pearson and Johnson (1978) propose the use of word association tasks in instructional settings to yield a diagnosis of what students already know and what they need to know about a key concept. The teacher throws out to a class a key concept or key word and asks the students to volunteer their associations on the word. Initial associations made by students may be of different types (superordinates, subordinates, attributes, definitions, synonyms, antonyms, contradictories, contraries, reverses, personal experiences, or even similar-sounding words). As students volunteer these associations the teacher writes them on the blackboard or overhead projector. The teacher and the students may go even further and organize the associations into a semantic map (Johnson & Pearson, 1978).

Reflection on these associations may then form the basis for further class discussion. Langer (1981) has found that through such class discussion students may significantly enrich their networks of associations and their vocabularies. In attempting to get students to “stretch” their concepts, Pearson and Johnson (1978) and Pearson and Spiro (1982) encourage the teacher to use analogies, comparisons, even metaphors to build bridges between what the students already know about a concept and what they may need to know in order to read and understand a particular text. Obviously, it is also helpful for the teacher to offer several examples of the new concept, as well as several examples of what it is not, so students develop a sense of the boundaries of the concept.

Such semantic mapping can be used jointly as both a pre- and post-reading activity. Johnson, Pittelman, and Heimlich (1986), and Heimlich and Pittelman (1986) claim that using a semantic map as a post-reading, follow-up activity “affords students the opportunity to recall, organize, and graphically represent the pertinent information read” (1986, p. 781). It further provides students with an opportunity to compare information gained from reading to information they possessed prior to reading, and to integrate new information.

Several organized approaches and methods have been proposed

language education we are, and should be, concerned with approaches that can improve the reading skills of all learners. In this paper I have tried to show how recent second language reading research conducted within the framework of schema theory may contribute toward the goal of improving students' foreign or second language reading skills.

**Patricia L. Carrell** is currently Dean of the Graduate School and Professor of English at the University of Akron in Akron, Ohio, U.S.A., to which she recently moved after 21 years at Southern Illinois University at Carbondale as Professor of Linguistics and Psychology. Her research on reading in a second language, which has been widely published in such journals as *TESOL Quarterly*, *Language Learning*, *Reading in a Foreign Language*, and *Modern Language Journal*, won her the Paul Pimsleur award for research in foreign language education from the American Council of Teachers of Foreign Languages (ACTFL) in 1985. She is the co-editor of two recent books on second or foreign language reading: *Interactive Approaches to Second Language Reading* (Cambridge University Press, 1988), and *Research in Reading in English as a Second Language* (TESOL, 1987).

### References

- Adams, M.J., & Collins, A. (1979). A schema-theoretic view of reading. In R.O. Freedle (Ed.), *New directions in discourse processing* (pp. 1-22). Norwood, NJ: Ablex.
- Alderson, J.C., & Urquhart, A.H. (1985). This test is unfair: I'm not an economist. In P. Hauptman, R. LeBlanc, & M. Bingham Wesche (Eds.), *Second language performance testing* (pp. 25-43). Ottawa: University of Ottawa Press. Reprinted in P. Carrell, J. Devine, & D. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 168-182). New York: Cambridge University Press, 1988.
- Anderson, R.C. (1977). The notion of schemata and the educational enterprise: General discussion of the conference. In R.C. Anderson, R.J. Spiro, & W.E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 415-431). Hillsdale, NJ; Erlbaum.
- Anderson, R.C., Reynolds, R.E., Schallert, D.L., & Goetz, E.T. (1977). Frameworks for comprehending discourse. *American Educational Research Journal*, 14, 367-381.
- Anderson, T.H. (1978). *Study skills and learning strategies*. Technical Report No.

require students to diagram the ideas and their inter-relationships within the text. Geva (1980, 1983), Taylor and Beach (1984), and Bartlett (1978) have all shown significant effects of teaching text structure in instructional settings with native speakers of English. Carrell (1985), as previously mentioned, showed significant effects of teaching Meyer's top-level organization patterns to readers of ESL. If students are not familiar with these rhetorical structures, such teaching serves not only to build the appropriate formal schemata, but also teaches students to activate the schemata. Mikulecky (1985) has developed classroom exercises which can be used to teach students to identify various top-level organizational patterns. Nuttall (1982) is also a good source of text-mapping classroom strategies and exercises.

Writing teachers have been aware for some time of the different rhetorical structures and the need to teach them explicitly to EFL/ESL students for composition. Now we are discovering that explicit awareness of these rhetorical structures, along with strategies for putting that knowledge to use in reading, can facilitate reading comprehension. Therefore, some EFL/ESL reading texts are borrowing rhetorical approaches quite directly from writing texts. One such text is entitled *Reading and Study Skills: A Rhetorical Approach*, by Kimmelman et al. (1984). That text departs from traditional approaches to teaching comprehension skills by grouping the skills within the framework of particular rhetorical modes and, as a result, making students aware of skills appropriate to these rhetorical modes—narrative, descriptive, analytical, and argumentative. Reading teachers may want to consult writing texts as sources for other ideas for teaching reading comprehension skills within the framework of rhetorical structure and organization.

## 6. Conclusion

Finally, I would like to underscore the importance of improving the reading skills of second and foreign language students. Effective and efficient second language reading skills are critical for many EFL/ESL students, especially for those whose primary source of language input may be via the written medium, as well as for those at advanced levels of proficiency, and for those with a need for English for academic purposes. As professionals in second

- Cambridge University Press.
- Floyd, P., & Carrell, P.L. (1987). Effects on ESL reading of teaching cultural content schemata. *Language Learning*, 37, 89-108.
- Geva, E. (1980). *Metatextual notions and reading comprehension*. Unpublished doctoral dissertation, University of Toronto.
- Geva, E. (1983). Facilitating reading comprehension through flowcharting. *Reading Research Quarterly*, 18, 384-405.
- Grabe, W. (1986). The transition from theory to practice in teaching reading. In F. Dubin, D.E. Eskey, & W. Grabe (Eds.), *Teaching second language reading for academic purposes* (pp. 25-48). Reading, MA: Addison-Wesley.
- Hall, M.A. (1981). *Teaching reading as a language experience*. Columbus, OH: Merrill.
- Haus, G.J., & Levine, M.G. (1985). The effect of background knowledge on the reading comprehension of second language learners. *Foreign Language Annals*, 18, 391-397.
- Heimlich, J.E., & Pittelman, S.D. (1986). *Semantic mapping: Classroom applications*. Newark, DE: International Reading Association.
- Hinds, J.L. (1983a, March). Retention of information using a Japanese style of presentation. Paper presented at the annual TESOL convention, Toronto, Canada.
- Hinds, J.L. (1983b). Contrastive rhetoric: Japanese and English. *Text*, 3, 183-195.
- Johnson, D.D., & Pearson, P.D. (1978). *Teaching reading vocabulary*. New York: Holt, Rinehart & Winston.
- Johnson, D.D., Pittelman, S.D., & Heimlich, J.E. (1986). Semantic mapping. *The Reading Teacher*, 39, 778-783.
- Johnson, N.S., & Mandler, J.M. (1980). A tale of two structures: Underlying and surface forms in stories. *Poetics*, 9, 51-86.
- Johnson, P. (1981). Effects on reading comprehension of language complexity and cultural background of a text. *TESOL Quarterly*, 15, 169-181.
- Kimmelman, J., Krantz, H., Martin, C., & Seltzer, S. (1984). *Reading and study skills: A rhetorical approach*. New York: Macmillan.
- Langer, J.A. (1980). Facilitating text processing: The elaboration of prior knowledge. In J.A. Langer & M.T. Smith-Burke (Eds.), *Reader meets author/bridging the gap* (pp. 149-162). Newark, DE: International Reading Association.
- Langer, J.A. (1981). From theory to practice: A prereading plan. *Journal of Reading*, 25, 152-156.
- Mandler, J.M. (1978). A code in the node: The use of a story schema in retrieval. *Discourse Processes*, 1, 14-35.
- Mandler, J.M., & Johnson, N.S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, 9, 111-151.
- Mandler, J.M., Scribner, S., Cole, M., & DeForest, M. (1980). Cross-cultural invariance in story recall. *Child Development*, 51, 19-26.
- Meyer, B.J.F. (1975). *The organization of prose and its effects on memory*. Amsterdam: North Holland.
- Meyer, B.J.F., & Freedle, R.O. (1984). Effects of discourse type on recall. *American*

- Educational Research Journal*, 21, 121-143.
- Mikulecky, B.S. (1985). Reading skills instruction in ESL. In P. Larson, E.L. Judd, & D.S. Messerschmitt (Eds.), *On TESOL '84* (pp. 261-277). Washington, DC: TESOL.
- Mohammed, M.A.H., & Swales, J.M. (1984). Factors affecting the successful reading of technical instructions. *Reading in a Foreign Language*, 2, 206-217.
- Nuttall, C. (1982). *Teaching reading skills in a foreign language*. London: Heinemann Educational Books.
- Pearson, P.D., & Johnson, D.D. (1978). *Teaching reading comprehension*. New York: Holt, Rinehart & Winston.
- Pearson, P.D., & Spiro, R.J. (1982). The new buzz word in reading is "schema." *Instructor*, May 1982,, 46-48.
- Rigg, P. (1981). Beginning to read in English the LEA way. In C.W. Twyford, W. Diehl, & K. Feathers (Eds.), *Reading English as a second language: Moving away from theory* (pp. 81-90). Bloomington: Indiana University Press.
- Rubinson, R.P. (1941). *Effective study*. New York: Harper & Row.
- Rumelhart, D.E. (1977). Understanding and summarizing brief stories. In D. LaBerge & S.J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension* (pp. 265-303). Hillsdale, NJ: Erlbaum.
- Rumelhart, D.E. (1980). Schemata: The building blocks of cognition. In R.J. Spiro, B.C. Bruce, & W.E. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33-58). Hillsdale, NJ: Erlbaum.
- Rumelhart, D.E., & Ortony A. (1977). The representation of knowledge in memory. In R.C. Anderson, R.J. Spiro, & W.E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 99-135). Hillsdale, NJ: Erlbaum.
- Schank, R.G., & Abelson, R.P. (1977). *Scripts, plans, goals and understanding*. Hillsdale, NJ: Erlbaum.
- Smith-Burke, M.T. (1980). Extending concepts through language activities. In J.A. Langer & M.T. Smith-Burke (Eds.), *Reader meets author | bridging the gap* (pp. 163-179). Newark, DE: International Reading Association.
- Spiro, R.J. (1978, December). *Beyond schema availability*. Paper presented at the Annual Meeting of the National Reading Conference, St. Petersburg, FL.
- Spiro, R.J. (1979). Etiology of reading comprehension style. In M.L. Kamil & A.J. Moe (Eds.), *Reading research: Studies and applications* (pp. 118-122). Clemson, SC: National Reading Conference.
- Stauffer, R.G. (1980). *The language experience approach to the teaching of reading*. New York: Harper & Row.
- Steffensen, M.S. (1988). The dialog journal: A method for improving cross-cultural reading comprehension. *Reading in a Foreign Language*, 5, 193-203.
- Steffensen, M.S., Joag-dev, C., & Anderson, R.C. (1979). A cross-cultural perspective on reading comprehension. *Reading Research Quarterly*, 15, 10-29.
- Stevens, K. (1982). Can we improve reading by teaching background information? *Journal of Reading*, 25, 326-329.
- Stroller, F. (1986). Reading lab: Developing low level reading skills. In F. Dubin, D.E. Eskey, & W. Grabe (Eds.), *Teaching second language reading for academic purposes* (pp. 51-76). Reading, MA: Addison-Wesley.

- Strother, J.B., & Ulijn, J.M. (1987). Does syntactic rewriting affect English for science and technology (EST) text comprehension? In J. Devine, P.L. Carrell, & D.E. Eskey (Eds.), *Research in reading in English as a second language* (pp. 91-101). Washington, DC: TESOL.
- Taylor, B.M., & Beach, R.W. (1984). The effects of text structure instruction on middle grade students' comprehension and production of expository text. *Reading Research Quarterly*, 19, 134-146.
- Urquhart, A.H. (1984). The effect of rhetorical ordering on readability. In J.C. Aldeson & A.H. Urquhart (Eds.), *Reading in a foreign language* (pp. 160-175). London: Longman.
- Widdowson, H.G. (1979). *Explorations in applied linguistics*. New York: Oxford University Press.

## Japanese-English Code Switching in Bilingual Children

Sandra S. Fotos

*Aoyama Gakuin Women's Junior College*

The results of an exploratory study of Japanese-English code switching in four bilingual children are presented here. Functional aspects of code switching are examined, and the results are consistent with findings reported in previous studies primarily involving Spanish-English switches. However, the two major linguistic constraints on code switching developed from consideration of switches between syntactically similar languages do not appear to hold in a number of Japanese-English switches. Furthermore, quantitative analysis of the switches reveals additional differences from earlier reports, these also linked to the issue of switching between syntactically dissimilar languages. The use of code switching as an enrichment device and an intensifying strategy to increase the dramatic effects of natural speech narratives is also considered.

### バイリンガル児童にみられる英語—日本語間のコード・スイッチングに関する研究

以下は、4人のバイリンガル児童の英語—日本語間のコード・スイッチングに関する研究をまとめたものである。この研究で得た結果は、これまで発表されてきたスペイン語—英語間のコード・スイッチング研究結果と共通部分が多い。しかしながら、ここでは、英語—日本語間では、類似の言語構造の場合に受けるといわれているコード・スイッチング上の制約は受けないのではないかとおもわれる事例について論じている。また、言語構造が異なる言語間のコード・スイッチングに関して過去に発表された結果と異なる結果についても論じている。さらに、子供が仲間同士で出来事や体験を語るとき、豊かな言語表現として、あるいは、ドラマティックな効果を上げるための強調の手段として用いるスイッチングについても論じている。

### 1. Introduction

Bilingual individuals often employ the grammar and lexicon of both their languages in a single utterance. This phenomenon,

known as code switching, has been examined since the 1950s, when general descriptive studies of multilingualism were published (see Heller, 1988, for a treatment of the early literature). By the 1960s a series of studies of code switching existed, but in many of these the alternation of languages was perceived as random or, worse, an indication of the lack of sufficient proficiency to converse fluently in either language (cf. the review in Martin-Jones & Romaine, 1986). However, a number of studies in the 1970s, particularly the work of Gumperz and his associates (1976), Pfaff (1979), and Valdes (1976), clearly demonstrated that language alternation was an important strategy for conveying social information such as the definition of role relationships between speakers and the establishment of feelings of solidarity and group membership. In addition, code switching was shown to function as a personal communication device, enabling bilinguals to organize and enrich discourse (Koike, 1987; McClure, 1981; Valdes, 1981).

At the same time as the sociolinguistic functions of code switching were being identified, studies of the linguistic processes involved revealed that far from being random, code switching was actually tightly governed by a number of linguistic factors, and a series of functional and syntactic constraints were demonstrated to exist in analyses of switches between Spanish and English (Pfaff, 1979; Poplack, 1980, 1981).

The current literature continues to reflect these two general directions of interest, as researchers investigate both the social and linguistic environments of switching in a number of languages, as well as in Spanish-English alternations (Azuma, 1987; Boeschoten & Verhoeven, 1987; Dabene & Billiez, 1986; Koike, 1987; Vihman, 1985). A brief review of some recent findings pertaining to this preliminary study of Japanese-English code switching in bilingual children is presented below.

### *1.1 Studies of The Social Functions and Linguistic Characteristics of Code Switching*

Language selection and alternation by bilinguals has been shown to be determined by situational factors such as the participants in the speech event and the type of discourse (Dabene & Billiez, 1986; McClure, 1981; McClure & McClure, 1988). This is called situational code switching. A second type is conversational

or stylistic code switching (Auer, 1988; Gumperz, 1976; Scotton, 1988; Valdes, 1976). Here code switching is functionally categorized as occurring for purposes of (a) quotation, where reported speech is code switched; (b) address specification, where specific individuals or groups are addressed or referred to by code switching; (c) emphasis, where statements made in one language are switched and repeated in the second language; (d) clarification, where a confusing statement is elaborated in the second language; (e) attention getting or retention, where discourse markers such as exclamations or interjections are code switched, (f) personalization versus objectivization, where one language is used to refer to factual events and the second language denotes the personal feelings of the narrator; and (g) topic shift, where certain topics tend to be discussed only in one language. Studies such as McClure's (1981) have examined the use of code switching for these functions in a variety of speech acts, and this report will examine instances of Japanese-English switches according to the above seven categories.

In addition, linguistic research, such as that conducted by Poplack (1980, 1981) on Spanish-English alternation, has suggested two general constraints determining where a code switch can take place: the free morpheme constraint and the equivalence constraint. The former says that code switching will not occur between a lexical form and a bound morpheme, such as *-ing*, unless the form has been integrated into the language of the morpheme. The latter says that the same type of word order before and after a switch must exist in both languages in order for a switch to maintain grammaticality. However, although these constraints were found in a number of studies to hold for Spanish-English switches, the limited research examining alternation between syntactically dissimilar languages, such as Japanese-English (Azuma, 1987; Loschky, 1989; Nishimura, 1985) and Turkish-Dutch (Boeschoten & Verhoeven, 1987) found switching behavior which could not be explained by these rules. Instances of Japanese-English code switching reported in this study will also be examined for operation of the two constraint rules.

Both linguistic and social function studies of code switching have found differences in the type of switching in relation to linguistic proficiency. In two studies, Poplack (1980, 1981) quanti-

tatively and qualitatively analyzed the speech of bilingual Puerto Ricans of different proficiency levels and divided switching into three main types according to structural features: intimate, emblematic, and noun switches. Intimate code switching was characterized as intra-sentential and was regarded as requiring considerable linguistic proficiency in both languages, since the code-switched item had to conform to shared underlying syntactic rules. Emblematic code switching referred to items which did not affect the grammaticality of an utterance, such as tags, frozen forms, and idiomatic expressions. This type of code switching is structurally less integrated into the discourse and was considered to require considerably less language proficiency. Noun switches, defined as the presence of an L2 noun in an otherwise L1 sentence, were found to be the most frequent in a number of studies (reviewed and continued in Poplack, 1981).

Several general results from these two studies have relevance for the data to be presented here: (a) switching tended to occur between major constituents such as sentences and clauses; (b) noun switches, however, were the most common; (c) there were no ungrammatical combinations of L1 & L2, regardless of proficiency level; (d) non-fluent bilinguals were able to code-switch frequently and maintain grammaticality by using emblematic code switching.

Similar results for Spanish-English intra-sentential switches were reported by Pfaff (1979), who suggested that such switches illustrate competence in the syntactic rules of both languages. Additional findings concerned structural triggers for switches. Pfaff reported that although most longer switches appear to be induced by functional considerations, some appear to be triggered by "code mixes." The need to define this term introduces the issue of nomenclature.

### *1.2 Variations in Terminology for Language Alternation*

There is a notable lack of agreement among researchers on a terminology for the different types of switching encountered in the literature. Traditionally, a distinction has been made between borrowing and switching. Researchers such as Poplack (1981) and Pfaff (1979) view a switch according to the degree of adaptation of the given item to the other language, one extreme being the complete phonological and morphological adaptation of the item to

the second language, this being called "borrowing," and the other extreme being the lack of any adaptation of forms to the second language, this being called "code switching." McClure (1981) used the term "code switching" to include both "code mixing" and "code changing," these terms being defined both socially and syntactically. "Code mixing" refers to the process whereby a single item (or items) from a second language is used in the first language because the item is either unknown in the language or is more appropriate for the social context of the utterance. "Code mixing" thus results in sentences composed primarily of the first language. This concept has also been referred to as "language mixing" (Boeschoten & Verhoeven, 1987; Pfaff, 1979). "Code changing," on the other hand, is considered to be a stylistic strategy used in social interaction, and takes place between constituent boundaries, often producing sentences that are sequentially L1 and L2.

Following Dabene & Billiez (1986), this study will use the general term "code switching" to refer to the communicative strategy of bilingual speakers which consists of alternating units of variable length belonging to two languages within a single turn. By this definition, code switching occurs every time the speaker switches from the language initially spoken to another. Thus, communicative units of all lengths are included in the analysis presented in this report.

### *1.3 Studies of Social and Linguistic Functions of Code Switching in Bilingual Children*

Up to this point, the problems discussed have derived from studies of bilingual adults. What of bilingual children? A pioneering study of the discourse function and linguistic constraints in the code switching of a number of Spanish-English bilingual children (McClure, 1981) found that code switching was largely a function of the participants rather than the setting or the discourse type. As with adults, conversational code switching took place for the seven functional categories discussed previously. Switches of single lexical items, considered by McClure to occur when the child either lacked the term in the L1, or found the L2 term easier to access, were common, and were often found to trigger longer code shift sequences. As was noted for adults, children who were fluent bilinguals, usually the older children, tended to code switch mainly

at the level of sentences or main clauses. Such shifts resulted in a sentence or independent clause in one language followed by a sentence or clause in the second language. Less fluent bilinguals avoided shifts at the intra-sentential level, preferring to shift at the word level instead, resulting in L1 sentences containing limited lexical items in the L2. As was found for adults, the degree of control of the two languages correlated with the type of code switching performed.

These general findings were confirmed in other studies of bilingual children reviewed in Grosjean (1982), who also noted that different types of code switching tend to occur depending on the age of the child. Children under eight years of age were found mainly to switch single lexical items, whereas in older children, inter- and intra-sentential code switching at both constituent and lexical levels was often used as a social marker of group membership and solidarity, as well as for functions such as emphasis, elaboration, and topicalization (Appel & Muysken, 1987; Grosjean, 1982), similar to its role in adult discourse.

Some confirmation of these results comes from two studies of very young children, two years of age or less, (Huerta-Macias, 1981; Vihman, 1985). Here code switching was observed to occur primarily with nouns and adjectives. An additional study of four-to-seven-year-old Turkish-Dutch bilingual children found that single lexical items were most commonly switched, these being nouns, adjectives, and, to a lesser extent, verbs. It was suggested that such switches expressed concepts acquired during secondary socialization which were not immediately accessible in the children's first language. Topic was found to be significant since, it was suggested, some topics are better discussed in one language than the other, due to social context or because of lexical shortcomings. The functional features of switches were consistent with those reported by McClure and resembled adult switches. Linguistically, the free morpheme constraint was not observed, and the operation of the equivalence constraint was questioned.

Additional studies of adolescents and younger children who were members of bilingual communities in France (Dabene & Billiez, 1986) and Italy (Auer, 1988) noted the use of code switching for different discourse functions, similar to patterns already observed in children and adults.

#### *1.4 Discourse Functions of Code Switching in Natural Speech Narratives*

A final aspect to consider in regard to the social and linguistic functions of code switching in the speech of bilinguals is the research findings on the use of code switching to structure and dramatize natural speech narratives. Three analytical methods have been used for examination of the structure of conversational narratives (Olsen-Fulero, 1986): Labov's "high point" analysis (1972), where the narrative is divided into components; episodic analysis, which focuses on the story grammar and the intentions and goals of the narrator; and dependency analysis, which deals with syntactic structure independent of the narrative content. Following the procedure used by Koike (1987) in her examination of code switching as a structuring and enrichment device in Spanish-English narratives, this study will also employ the structural categories used by Labov (1972) in his work on narratives in Black English Vernacular. Here, a narrative is divided into the following components: (a) an abstract, which is a short summary of the story; (b) an orientation, which specifies the time, place, situation, and characters; (c) complicating actions, which are the main events of the narrative; (d) an evaluation, which tells the point of the narrative; (e) a resolution, which tells the outcome of the narrative; and (f) a coda, which ends the narrative and brings the participants back to present time. Both Koike (1987) and Valdes (1981) have noted the presence of code switching at the component boundaries of narratives related by bilingual adults, and have suggested that such switching serves the functions of delineating critical parts, holding the listener's attention, and moving the action forward.

Code switching as a discourse device to enrich Spanish-English narratives has also been explored in studies by Genishi (1981) and Koike (1987). Using Labov's narrative enrichment devices (1972) as the units of analysis, switches have been shown to function as intensifiers, which select and intensify an event; comparators, which compare events that did occur to those that did not; correlatives, which bring together two events that occurred so that they are conjoined in a single clause; and explicatives, which describe events not familiar to the listener. Switches have been shown to be used (Koike, 1987) for dramatic effect, topicalization, and the

introduction of new information at different points in the narrative.

In addition to the structural and functional characteristics of Japanese-English code switching, the dramatization functions of code switching in a bilingual child's natural speech narrative will also be examined in the following report.

## 2. Subjects and Methods

Data with a running time of four hours were collected from four Japanese-English bilingual children attending an English-language international school in Tokyo, Japan. Two sister-brother sets were used. At the time of the study, the sisters were eleven years old, and in the fifth grade, and the brothers were seven years old, and in the first grade. One sibling set is from a culturally mixed family, with a Japanese mother and an American father. The parents of the second set are both American but the children attended Japanese National Day Care Centers for four years before transferring to the international school and have maintained their proficiency in Japanese through after-school lessons in reading and writing from a native speaker.

Data were collected on two separate occasions by leaving a tape recorder running in the rooms where the children were playing. The children were told only that the author wished to record their speech. One additional recording was made while all four children were eating snacks together after school. In this exploratory study, only 40 minutes of the data are presented for analysis: three narratives and one conversation for the sisters, and three narratives and one conversation for the brothers.

For the quantitative analysis of the code switching data, raw frequencies were tabulated and percents calculated for the number of switches, defined as each time there was a language change regardless of the length of the utterance, and the types of items switched. The Chi-square test was used as a statistical measure of the significance of different shifts, and the alpha level was set at .05. One-way Chi-square tests with 1 degree of freedom were corrected for continuity following Hatch & Farhady (1982). In addition, the linguistic environment of the switches was examined to see whether the constraint rules previously discussed applied in to code switching between two syntactically dissimilar languages.

For the functional analysis, the categories proposed in the literature (code switching to show quotation, to clarify meaning, to get and hold attention, to change the topic of discussion, to use special nouns culturally linked to one of the languages, and to personalize or objectivize events) were used to examine switching produced by the four children. For the narrative analysis, due to limitations of space and the desire to present the entire protocol, only the longest conversational narrative, related by one of the sisters, was examined in terms of the Labovian narrative components and enrichment devices previously defined. However, all narratives investigated were similar in structure and in code switching function. Following Koike (1987), the language used for each section was investigated and code switched items were inspected to determine their function in setting off parts of the narrative—building dramatic effect through the contrast of languages.

Transcription of the data is in standard English orthography, with glosses reported in the Results and Discussion section, following the procedures of Cziko & Koda (1986). Japanese is translated first morpheme by morpheme, then by sentence. Coding of the data is consistent with syntactic categories used in Poplack (1980). The author was able to consult both Japanese and English colleagues on the translation and categorization of questionable items, so inter-rater opinions were available in some cases. The children were also consulted during the coding and glossing procedures, and Kuno (1973) was used as the primary reference for Japanese grammar points.

### 3. Results and Discussion

The respective results and discussion of the quantitative analysis, the functional analysis, and the narrative analysis will be presented separately.

#### *3.1 Quantitative Analysis of Code Switching and the Constraint Rules*

Table 1 shows the breakdown of language use and code switching percentages for the older and younger children. The unit of analysis is the C-unit, or Communicative unit (Crookes, 1988), which is similar to a T-unit, and consists of a main clause plus

**Table 1**  
 Frequencies and Percentages of Code Switching in Both Languages and Between and Within Sentences for Older and Younger Children

	11 year olds	6-7 year olds
<u>Total number of C. S.</u> (n=153)	93	60
Japanese to English	46 (49%)	31 (52%)
English to Japanese	47 (51%)	29 (48%)
<u>Total intra-sentential C. S.</u>	48 (52%)	39 (65%)
<u>Total inter-sentential C. S.</u>	45 (48%)	21 (35%)
<u>Intra-sentential C. S.</u> (n=87)		
Japanese to English	23 (49%)	19 (48%)
English to Japanese	25 (51%)	20 (52%)
<u>Inter-sentential C. S.</u> (n=66)		
Japanese to English	22 (48%)	12 (57%)
English to Japanese	23 (52%)	9 (43%)

subordinate clauses or isolated phrases not accompanied by a verb but which have communicative value. Switches were considered to occur whenever there was a language change, and a total of 153 switches were counted in the data discussed here.

Table 2 reports the percentages of switches made in the different syntactic categories for total switches, switches from Japanese to English, and switches from English to Japanese. Since the percentages were very similar for both age groups, the data are combined.

No significant difference was observed between the number of switches from Japanese to English and the number of switches from English to Japanese ( $\chi^2 = .42$ ,  $df=1$ , n.s.). Of the total switches, 44% consisted of single lexical items such as nouns (16% of the total switches) conjunctions (14%) and fillers (14%); furthermore, 16% of the total were switches of entire sentences. These two types of switches, either of single items or of entire sentences, comprised a significant number of the total, 107 switches out of 153 ( $\chi^2 = 23.53$ ,  $df=1$  ;  $p < .05$ ). According to the literature on Spanish-English

**Table 2**  
Percentages of Switches in Different Syntactic Categories

	Total	Japanese- English	English- Japanese
	(n=153)	(n=81)	(n=72)
<u>Intra-Sentential C. S.</u>	(%)	(%)	(%)
noun:	16	13	19
independent clause:	4	6	5
adjective:	0.6	-	1
dependent clause:	5	10	1
adverb:	7	6	6
verb:	6	4	8
possessive:	1	-	2
<i>wa /ga</i> topic marker:	2	-	5
conjunction:	14	2	7
subject+topic marker:	0.6	-	1
<u>Inter-Sentential C. S.</u>	(%)	(%)	(%)
sentence:	16	13	12
interjection:	2	1	2
filler:	14	9	-
tag:	4	1	8
quotation:	3	-	8
exclamation	6	4	10
back channel agree:	4.4	6	5

\*for percentage figures, numbers below 1.0 are not rounded to the next number

switches, such a syntactic switching pattern is more indicative of lower proficiencies, although for all levels of proficiency, noun switches were found to be most common.

Inspection of 153 switches, 93 switches made by the older children and 60 by the younger children, suggests that there were few differences between the two age sets in switching pattern expressed in percentages, except that the younger children tended to switch more intra-sententially, 65% compared to 52%. This difference was not significant (frequencies weighted out of 50 ;  $\chi^2=$

.55,  $df=1$ , n.s.). However, because of the small number of switches analyzed and the fact that the children were sibling sets, consideration of the lack of significant differences in observed switching behavior must be set aside, and the issue of age-related switching cannot be addressed in the present study.

Addressing the issue of the language proficiency of the children in this study, results on the English language achievement portion of the Iowa Test, a standardized achievement test used in the U. S., placed the older children in the 87th and 96th percentiles respectively, indicating high levels of English proficiency. In addition, on the basis of proficiency examinations and previous achievement tests, both girls were placed in the native speaker Japanese language ability group. Iowa test results for the two boys were not available, but their English reading and language skill test grades were above the class average, and they were also placed in the native speaker Japanese language ability group. Therefore, it would not be correct to interpret the code switching results presented here as characteristic of so-called "semilingual" children, who are said to lack skills in either language.

In conjunction with this latter point, it must be noted that research on code switching has often been carried out among immigrants whose L1 is perceived by the host community to be of lower economic and social status than the national language, the L2 of the research subjects. This has been true for a number of the Spanish-English code switching studies (reviewed and continued in Poplack, 1981) and has been reported for studies of Turkish immigrant children in the Netherlands (Boeschoten & Verhoeven, 1987) and for children of Spanish, Portuguese, and Algerian immigrant communities in France (Dabene & Billiez, 1986). However, the status of English in Japan is somewhat different. English holds a respected position, so much so that it is a significant component of the important high school and university entrance examinations. Furthermore, although the children in the present study attended an international school where English was the medium of education, Japanese language skills were also socially reinforced. Thus, in contrast to a number of other studies, it can be suggested that social factors appeared to support, rather than operate against, the development of proficiency in both languages in the subjects studied here.



**it's *takai* than most places**

expensive

This last utterance is one of the few ungrammatical forms recorded.

It should be noted that some linguists might argue that Japanese particles are not bound in the sense that bound morphemes exist in English or Spanish, and if this point is accepted as valid, the constraint rule is not violated in the data presented above. However, it is beyond the scope of this paper to enter into such a debate, and it is suggested that use of Japanese topic markers in English sentences attached to English topics represents a compromise of the constraint rule. Clearer violations of the constraint rule may be seen below.

English morphemes were frequently attached to Japanese nouns, making them plural or genitive: for example, *mimizu-s* ('worm' plus English plural *s*). Japanese verbs were often marked with English tense morphemes. In the next example the English past tense suffix '-ed' was attached to a Japanese verb base code switched into an English sentence.

(3) **and then she got *yukai*-ed.**

kidnap

Loschky (1989) also noted cases in her data that violated the free morpheme constraint (Poplack, 1980), a few even involving switching taking place within a single word.

The equivalence constraint, completely rejected by Azuma (1987), especially when Japanese coordinating conjunctions were used, was compromised by only a few utterances in this study, as most switches avoided intra-sentential constituents.

(4) ***konnani chisakatta kara***  
 this much small+past since  
**so I got that**

since it was this small

Example 4, one of the few instances of an intra-sentential switch involving dependent clauses, illustrates the difficulty of maintaining grammaticality. Here, the messy problem of the Japanese SOV word order is avoided since the verb is in English, and the equivalence constraint is upheld. A much more common strategy, as noted, is switching involving independent clauses:

- (5) **You put your hands like this**  
**me o tojite** close your eyes  
 eye object marker close + imperative  
**and sit there**

### 3.2 *Functional Analysis*

Since the setting and the participants were constant during the recording of the discourse examined here, these variables of situational code switching are not addressed. However, the topic appears to have been important in determining language changes since many nouns were culturally and linguistically linked to particular topics or activities. Although it would have been desirable to establish the links between particular noun phrases referring to encoded topics and the language preferred for them prior to and independent of post hoc analysis, such a procedure was not possible in this exploratory study. Consequently, it can only be suggested that the following are examples of topic-linked switches. Example (6) presents a switch from previous speech in Japanese to English when discussing school, which is an English-only environment, using a restatement in Japanese for dramatic effect, and maintaining key nouns in English.

- (6) **my day was awful**  
**boku no day wa awful** my day was awful  
 I (masculine) possessive topic marker  
**datta no ne**  
 copula + past tense particle tag

A very common shift involved reference to money, nearly always made in Japanese:

- (7) **I saw this bracelet that was for *san byaku en*** 300 yen

The younger children played video games during part of the recording and many noun switches referred to specialized terms linked to this activity:

- (8) **but my *kougeki* is** attack  
**pretty *umai* too** skillful

Code shifting for quotation, a phenomenon widely reported in a number of other studies (McClure, 1981), was also present in several instances, where reported speech was framed in English, then delivered as a direct quote in Japanese.

(9) **and then he said**

*oshi* exclamation  
**that was close !**  
*ougon no ryuu da* it's the gold dragon!  
 golden possessive dragon is (emphatic)

Switching for emphasis and clarification, both common occurrences reported for the speech of Spanish-English bilingual children, often involve the code switched repetition of an utterance. This can be seen in Example 6 or in the following, where the thought is repeated three times, with increasing degrees of emphasis.

(10) **they were really fake**

**but they were exactly like Reeboks**  
*honmono mitai* they look like the real thing  
 real thing look like  
*zettai ni honmono ni mite iru* they totally look like the  
 real thing  
 totally particle real thing particle  
 look like stative

Code switched repetitions to clarify ambiguity were also observed:

(11) **the hammer was in the hand like this**

*kou iu fuu ni* like this  
 this say manner particle

Use of a code switch to emphasize a particular part of the utterance is referred to as topicalization, as in the following example:



**A Trip to Osaka**

1. and guess what
2. *kondo Osaka ni iku toki ni ne* when I go to Osaka next  
next Osaka to go time particle tag
3. um
4. we might borrow
5. an
6. um
7. a little little little room in um
8. the *nikai date* Shinkansen  
a bullet train with two levels
9. and you know
10. I . . . I got this thousand yen
11. from a
12. my grandma's friend
13. for *toshidama* New Year's gift of money
14. and that was the day when  
I took it to
15. um
16. when I was going to Osaka
17. *dakara* so
18. *ne* well
19. *dakara* so
20. uh I was going to use it all up
21. and I got this chapstick with a little color
22. you know
23. two for three eighty
24. *konnani chisakatta kara* since they were so small  
in this way small+past tense since
25. so I got that
26. *soshite* and
27. *Doraemon no buku katta no ne* I bought a Doraemon book  
Doraemon possessive book buy  
+ past tense particle tag
28. *soshitara* and then
29. *sore nanahyaku en desho* that was seven hundred yen,  
that seven hundred yen is question wasn't it
30. I had to get a present for that person
31. *dakara* so

32. *nihyaku en ni naru n* it came to two hundred yen  
two hundred yen particle become particle
33. *ah* uh no
34. *sanbyaku en agete ni* it added up to three hundred yen  
three hundred yen rise particle
35. *sono toki okane nakatta kara* since I didn't have any money  
that time money not+ past tense since then
36. *dakara ano. . .* so uh
37. *nandattake* what should I say
38. *ah* uh
39. *kashite moratta no ne* I got a loan  
loan receive+past tense particle tag
40. *obaachan ni* from grandma
41. *soshitara* and then
42. um
43. when I got to the hotel *ne* tag
44. um
45. I I saw this bracelet that was for  
*san byaku en* three hundred yen
46. and I
47. I wanted it
48. so my grandma goes *oke* ok
49. *ne* tag
50. *soshite* and
51. *soshitara* and then
52. *kattara* she bought it
53. *ano* uh
54. *kondo mata* again next time
55. *Hankyuu depaato ni ittara* when we went to the Hankyu  
Hankyu department store to go dept. store  
+past tense +when
56. um
57. you see
58. um
59. there was a ring I wanted
60. this one
61. and
62. it was um *sen sanbyaku en* one thousand three hundred yen
63. so I got this too

64. so it was like I only had  
 65. um  
 66. *sen en* one thousand yen  
 67. but I got  
 68. a lot of stuff for *sen en* one thousand yen  
 69. I didn't really get it for *sen en* but one thousand yen  
 70. it was like two thou. . .  
 71. I don't know  
 72. everything's pretty cheap there  
 73. but in Tokyo  
 74. *nan de takai no* why is it expensive  
     why expensive particle

The narrative begins in English. An abstract is lacking, but the first part of the orientation, line 2, is delivered in Japanese. The rest, lines 3-8, are said in English except for a unique referent, a special type of fast train. Six complicating actions, or main events of the narrative, can be identified: lines 9-16; lines 17-25; lines 26-34; lines 35-40; lines 41-52 and lines 53-63, these dealing with various purchases made during the trip to Osaka. It is interesting that many of these complicating actions begin with a Japanese coordinating conjunction such as *dakara* (so, because), *soshite* (then), or *shoshitara* (and then), even though further speech may be switched to English. This type of code switching at the boundaries of narrative sections has been reported previously as a common intensifying strategy, moving the action forward and holding attention (Koike, 1987). Comparators, which compare events which did occur to those which did not, were not observed. However, the use of a correlative, which considers events which might happen in relation to the event being presented, is seen in lines 4-8, when the narrator tells of her possible future trip to Osaka on a special train. In this case, no shift was observed other than the lexical shift previously reported.

Switching for quotation or paraphrase has already been noted, but the example in line 48, where the grandmother says "ok," is especially notable, since the English expression was phonologically integrated into Japanese morphology as a borrowing.

The evaluation and result, lines 64-70, is delivered in English, as is the first part of the coda, lines 71-73. However, the last part

of the coda is switched to Japanese. The other girl commented on this final statement with her own code switched Japanese coda, partially given in Example 2. The use of Japanese as the language of choice for the coda and the occurrence of a comment upon the final statement by the listener, the comment also being expressed in Japanese, was seen in several of the other narratives examined in this study.

A similar pattern has been noted in Spanish-English narratives (Koike, 1987) where the evaluative and resultative elements were in one language, and the complicating actions in the other. It was suggested that this is due to personalization versus objectification issues, where factual events are presented in one language and the narrator's personal feelings are presented in the other language. Evidence for the same type of dynamics can be seen in the orientation and in lines 26-31, where a statement of what was bought, how much it cost, and how much money remained was given in Japanese, the language of objectification, and the personal action of having to buy a present was given in English. In general, this narrative uses Japanese as the backgrounding language to recount actions performed, while English expresses the speaker's feelings of wanting to buy things and being successful. Perhaps this choice is due to the tendency of the Japanese language to omit the first person pronoun (Kuno, 1973), so that bilingual individuals may prefer to express personal feelings in English, where a sentence can freely begin with I.

#### 4. Conclusions

The results of this exploratory study of code switching in four Japanese-English bilingual children are in agreement with findings from previous studies primarily involving Spanish-English switches. As with most studies of code switching, nouns and other single lexical items were most frequently switched, and grammaticality was observed in nearly all utterances, regardless of the combination of languages. Conversational code switching took place for the discourse functions of quotation, clarification, emphasis, attention getting and retention, and topic shifts. Narrative analysis indicated that code switching occurred along narrative section boundaries as an intensifying strategy and was used for

dramatization and issues of personalization versus objectification.

Certain differences from the Spanish-English data were also noted. Switching tended to occur between sentences or independent clauses, rather than intra-sententially, and it was suggested that this is caused by the lack of syntactical symmetry between Japanese and English. In addition, as was found in previous studies of Japanese-English switching, the free morpheme restraint was not upheld.

To reiterate remarks made at the beginning of this paper, code switching is neither random nor an indication of the lack of linguistic proficiency. Rather, it is rule governed and systematic, with well-described discourse functions. The data presented in this study is consistent with a positive view of Japanese-English bilingual children who are developing proficiency in both languages and who skillfully use code switching to enrich and vitalize their speech. As the number of bilingual individuals increases in Japan, a greater awareness and understanding of this phenomenon is necessary, not only for educators and parents, but for the general public as well. It is hoped that further and more extensive studies of Japanese-English code switching will be undertaken in the near future to add to our knowledge and appreciation of this interesting facet of bilingualism.

*This article is a version of a paper presented at the Annual Convention of the Japan Association of Language Teachers, Okayama, Japan, November, 1989.*

**Sandra S. Fotos** is a lecturer in English at a women's junior college in Tokyo and an Ed.D. student in Temple University Tokyo's TESOL graduate program. Previously she taught and developed curriculum in California, Hawaii, Guam, Kenya, and the Philippines. Her current research interests involve the role of formal instruction in second language acquisition.

### References

- Auer, J. (1988). A conversation analytic approach to codeswitching and transfer. In M. Heller (Ed.), *Codeswitching: anthropological and sociolinguistic perspectives* (pp. 187-214) Berlin: Mouton de Gruyter.
- Appel, R., & Muysken, P. (1987). *Language contact and bilingualism* Baltimore,

MD: Edward Arnold.

- Azuma, S. (1987). Conversational code-switching in Japanese/English *JACET Bulletin*, 18, 1-20.
- Boeschoten, H., & Verhoeven, L. (1987). Language-mixing in children's speech: Dutch language use in Turkish discourse. *Language Learning*, 37, 191-215.
- Crookes, G. (1988) Structural discourse analysis units for second language research. *Working Papers in ESL*, 7, 143-167. Hawaii:University of Hawaii.
- Cziko, G., & Koda, K. (1987). A Japanese child's use of stative and punctual verbs. *Journal of Children's Language*, 14, 99-111.
- Dabene, L., & Billiez, J. (1986). Code-switching in the speech of adolescents born of immigrant parents. *Studies in Second Language Acquisition*, 8, 309-325.
- Genishi, C. (1981). Codeswitching in Chicano six-year-olds. In R. Duran (Ed.) *Latino language and communicative behavior*. (pp. 133-152). Norwood, NJ: Ablex
- Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism* (pp. 179-227). Cambridge, MA: Harvard University Press.
- Gumperz, J. (1976). The sociolinguistic significance of conversational code-switching. *Working Papers of the Language Behavior Research Laboratory*, 46, Berkeley, CA: University of California,
- Hatch, E. & Farhady, H. (1982). *Research design and statistics for applied linguistics*. Rowley, MA.: Newbury House.
- Heller, M. (Ed.) (1988). *Codeswitching: Anthropological and sociolinguistic perspectives*. Berlin: Mouton de Gruyter.
- Huerta-Macias, A. (1981). Codeswitching: All in the family. In R. Duran (Ed.), *Latino language and communicative behavior* (pp. 153-169). Norwood, NJ: Ablex.
- Koike, D. (1987). Code switching in the bilingual Chicano narrative. *Hispania*, 70, 148-154.
- Kuno, S. (1973). *The structure of Japanese language*. Cambridge, MA: Massachusetts Institute of Technology Press.
- Labov, W. (1972). *Language in the inner city*. Philadelphia, University of Pennsylvania Press.
- Loschky, M. (1989 November). A syntactic analysis of English-Japanese intrasentential code-switching. Paper presented at the 28th Annual Convention of the Japan Association of College English Teachers, Kyushu, Japan.
- Martin-Jones, M. & Romaine, S. (1986). Semilingualism: A half-baked theory of communicative competence. *Applied Linguistics*, 7, 26-38.
- McClure, E. (1981). Formal and functional aspects of the codeswitched discourse of bilingual children. In R. Duran (Ed.) *Latino language and communicative behavior* (pp. 69-94) Norwood, NJ: Ablex.
- McClure, E., & McClure, M. (1988). Macro- and micro-sociolinguistic dimensions of codeswitching in Vingard (Romania). In M. Heller (Ed.), *Codeswitching: Anthropological and sociolinguistic perspectives* (pp. 25-52). Berlin: Mouton de Gruyter.
- Nishimura, M. (1985). *Intrasentential codeswitching in Japanese and English*. Unpublished doctoral dissertation. University of Pennsylvania.

- Olsen-Fulero, L. (1986). A review of C. Peterson & A. McCabe (1983), *Developmental psycholinguistics: Three ways of looking at a child's narrative*. *Journal of Children's Language*, 13, 450-454.
- Pfaff, C. (1979). Constraints on language mixing: Intrasentential code-switching and borrowing in Spanish/English. *Language*, 55, 291-318.
- 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. (1981). Syntactic structure and social function of codeswitching. In R. Duran (Ed.), *Latino language and communicative behavior* (pp. 169-185). Norwood, NJ: Ablex.
- Scotton, C. (1988). Codeswitching as indexical of social negotiations. In M. Heller (Ed.), *Codeswitching: Anthropological and sociolinguistic perspectives* (pp. 77-96). Berlin: Mouton de Gruyter.
- Valdes, G. (1976). Social interaction and code-switching patterns: A case study of Spanish/English alternation. In G. Keller, R. Teschner, & S. Viera (Eds.), *Bilingualism in the bicentennial and beyond* (pp. 53-85). New York: Bilingual Press.
- Valdes, G. (1981). Codeswitching as deliberate verbal strategy: A microanalysis of direct and indirect requests among bilingual Chicano speakers. In R. Duran (Ed.), *Latino language and communicative behavior* (pp. 95-107). New Jersey, Ablex.
- Vihman, M. (1985). Language differentiation by the bilingual infant. *J. Child Language*, 12, 297-324.

**Pausology and Listening Comprehension:  
Theory, Research, and Practice**

**Roger Griffiths**

*Nagoya University of Commerce and Business Administration*

In this paper the investigation of temporal variables is presented as an example of the type of research on microprocesses which can be applied to actual language teaching. Such research is not derived from a strong theoretical network (which all the evidence suggests is unattainable) but it builds upon sound empirical work in L1 pausological studies and L2 input studies. A brief review of early L2 studies of temporal variables (mostly of NS-NNS speech rate) indicates them to be methodologically flawed to the extent that their findings must be discounted. A degree of optimism is, however, found in the improved quality of recent studies, which are reviewed in some detail. Particular attention is paid to studies which demonstrate that listening comprehension can be facilitated by the manipulation of temporal variables.

**1. Introduction**

In this paper I not only want to discuss the findings of a number of pausological investigations into listening comprehension, but also want to examine how they fit into the wider picture of research in applied linguistics. I would first like to make a number of observations on the status of theory generally in SLA before considering the status of theory on listening comprehension in particular. After that I will report on a number of studies which show that temporal variables can be modified to facilitate listening comprehension (throughout considered to be primarily comprehension of teacher-talk). Such research, minimal in theory, is then shown to have implications for practical application.

**2. Theory, Research, and Practice in L2 Learning**

Feyerabend (1970) has observed that:

The attempt to create knowledge needs guidance, it cannot start from nothing. More specifically it needs a theory, a point

of view that allows the researcher to separate the relevant from the irrelevant, and that tells him in what areas research will be most profitable. (p. 201)

Feyerabend is, however, specifically discussing the position of theory in the mature sciences where the proposition that "No information is collected in vacuo" (Medawar, 1984, p. 17) is commonly held (O'Hear, 1989, p. 89). This position is reflected in the following statement by Popper (1970):

A scientist engaged in a piece of research, say in physics, can attack his problem straight away. He can go at once to the heart of the matter: that is the heart of an organised structure. For the structure of scientific doctrines is already in existence; and with it a generally accepted problem situation. This is why he may leave it to others to fit his contribution into the framework of scientific knowledge. (p.51)

In terms of structures of knowledge in the mature sciences there is, as Popper puts it, "an edifice" (1970, p. 51). In the social sciences, however, even "well-shaped building stones are hard to come by" (Cronbach, 1987, p. 402), and there is no such edifice.

The major characteristic of a mature science has been conceptualized as it being governed by a single "paradigm," what Kuhn also calls a "disciplinary matrix" (1970a, p. 10). It is the lack of such a paradigm that most prevents the social sciences from qualifying as mature sciences. They are variously described, therefore, as "psuedosciences" (Feyerabend, 1970, p. 202), "spurious sciences" (Popper, 1970, p. 58), and "proto-science" (Kuhn, 1970b, p. 244). Adding to this list, Masterman (1970, p. 73) also identifies "non-paradigm science," which she describes as, "the state of affairs right at the beginning of the process of thinking about the state of the world, i.e., at the stage when there is no paradigm" (p. 73). She contrasts this with "multi-paradigm science", where she notes, "far from being no paradigm, there are on the contrary, too many" (p. 74).

Masterman considers the latter to reflect the overall situation in the psychological, social and information sciences, and she notes that at this stage, "discussion on fundamentals remains, and long-run progress (as opposed to local progress) fails to occur" (p. 74).

While it is easy enough, using the above criterion, to see that

applied linguistics does not qualify as a mature science, it is difficult to see where it fits into Masterman's conceptualization. Despite Brown's (1989, p. 169) recent baffling statement that one of the gratifying things about SLA research "is that it is research based on theory," it is doubtful that anything we have in SLA even qualifies as a "theory" (Gregg, 1984, p. 79), much less a paradigm. That, of course, would place the discipline in the first category. The total disagreement and seemingly endless debate over the legitimacy of fundamental assumptions and methods would indicate the second. It is, after all, not difficult to recognise something of our field in Feyerabend (1970, p. 202) and Kuhn's description of such endless debate: "The debates of 'pre-science' with their universal criticism and their universal proliferation of ideas are often directed as much to the members of other schools as...to nature."

The major problem at the stage in which "anything goes" (when no one paradigm marks the discipline) is that, as Krige puts it, "anything goes...means that, in practice, everything stays" (1980, p. 142). And much of what stays, does so despite incisive and relentless criticism, for example, that of Gregg (1984, 1986, 1988) and McLaughlin (1978, 1987) on Krashen's theory of SLA. Consideration of the contributions of others that have arrogated the name of "theory" in SLA, confirms this view (e.g., see Gregg, 1984, p. 96, on the theoretical contributions of Gattegno, Lozanov, and Curren). Ultimately, as O'Hear (1989) points out: "If a ruling theory does not fit the facts, someone in the end is going to blow the whistle if only because he will make his reputation by doing so" (p. 214). The facts are, however, seldom easily come by and not always scrupulously collected or reported (e.g., see Beretta on Krashen's reporting of empirical testing of Total Physical Response methods, 1986, p. 433).

It is clear, therefore, that none of what Ellis (1985, p. 248) generously calls a "plethora of theories" in SLA has resulted in anything like a common paradigm being adopted in research: no theoretical stance (for none of the so-called "theories" can be described as more than this) has, in fact, been a significant landmark in the progress of the discipline. However, research has gone on, with much of it being of the research-before-theory type (see Long, 1985a). While not wishing to espouse a naive inductivist position, it must be observed that much of the data gathered under

this strategy is clearly not “theory-laden.” The position is, in fact, much more akin to what O’Hear (1989) calls the “weak thesis” (p. 82) on this issue. This is perhaps worth quoting at length:

In considering the relationship between theory and observation in science, it is important to distinguish a weak thesis about the suffusing of observation by theory from a much stronger one. The weak thesis says that all observations are conditioned by presuppositions, assumptions regarding similarity and dissimilarity, directions of interest, and so... the fact that there are interests and schemes of classification behind any observation of the world does not amount to any elevated sense of theory. (pp. 82-83)

O’Hear also notes: “That all observation involves presuppositions and assumptions is undoubtedly true, but this does not imply that there is point in drawing distinctions between more and less theoretical levels of observation” (pp. 90-91). Because there is little agreement on the appropriateness of specific theories and because the theories we have are seldom explicit enough to indicate the direction of research programs, research in SLA conforms very much to O’Hear’s “weak thesis” on theory. Nor, to involve another dimension in the discussion, is there a strong relationship between explicit theory and practice in SLA. “Divorce” (Sinclair, 1989) does not seem to quite capture the quality of the relationship, as the two have never been closely attached (even the audiolingual approach bore only “a vague resemblance to an early version of Thorndikean association theory” [Carroll, 1966, p. 104]), but it is certainly the case that there is little intimacy. As Kasper (1988) recently wrote, “practitioners in the language teaching profession will be disappointed if they look for specific instructions on L2 teaching in the available literature...” (p. 5). But progress has been made, particularly in the area of input studies which have resulted in some practical applications. Long et al.’s (1984) work on “wait time” is an example of such intervention, although Long (1985a, 1985b) acknowledges that generally research on input is not derived from, nor has yet been incorporated into, current theories.

However, despite the absence of a unified theory, what research findings can do, as Lightbown (1985, p. 183) points out, is to

“answer some short-term questions about ‘what works.’” She continues, “Such short-term results are useful for getting teachers from day to day while they await the fulfillment of the great expectations of what might be called ‘basic’ or ‘non-applied’ (or even ‘pure’) research.” But getting from day to day is exactly what is required. This orientation is also congruent with Corder’s (1984, p. 58) view that “it is the task of the applied linguist to make practical use of whatever knowledge is available at the time.” In addition, the possibility of arriving at a total explanation must be extremely remote.

Cronbach (e.g., see 1975), is certainly of this view with regard to mainstream psychology, in which the quest for an overarching learning theory has been abandoned: in attempting to answer the question, “Should social science aspire to reduce behavior to laws?” (1975, p. 116), he looks back on the thirty years in which it has been attempted and laconically concludes, “I think most of us judge theoretical progress to have been disappointing” (p. 116). He further states: “The goal of our work...is not to amass generalizations atop which a theoretical tower can someday be erected.... The special task of the social scientist in each generation is to pin down the contemporary facts” (p. 126). To this one would only wish to add (even taking into account all the necessary caveats), “and use them.”

This certainly appears to be a realizable aspiration for the SLA researcher. It is a view which necessarily gives rise to the detailed investigations (described by Long et al. [1984, p. 3] as “research on microprocesses”) which are the hallmark of “normal science” (Kuhn, 1970a, p. 10) and which may have practical relevance and application. It is also the type of research referred to by Kasper (1988), who having noted that the large scale “methods” studies of the 60s and 70s failed largely as a result of failing to control the numerous variables involved, notes, “Recent classroom studies wisely examine more closely defined issues, e.g., features of teacher-talk...” (p. 5). Such an approach involves continuing to direct research resources towards the practice of one of the world’s great growth industries—classroom language teaching. Given, as Gregg (1988, p. 75) puts it, “the vast armies of unsuccessful L2 learners,” it is clear that it is not the outstanding quality of our product that

attracts consumers but a seemingly inexhaustible demand.

Without expecting nomothetic, or law-establishing, networks to emerge, it is assumed, however, that research should follow “the methods of the hard sciences” (Eysenck, 1986, p. 397). Eysenck’s view that psychology should consist of the “arduous work of providing actual proofs” (p. 397) is, in fact, central to the approach adopted in the research described here. There should, in short, be no room for choices over respective positions on academic issues in SLA to be made “on the grounds of personal preference, untrammelled by factual and general scientific consideration” (p. 397).

I see my own research on listening comprehension and pausology within this framework. However, before discussing pausological research, it is probably necessary to comment briefly on the current state of theory and research on comprehension generally and L2 listening comprehension in particular.

### **3. Comprehension: Theory and Research**

According to Ehrlich (1982), the term comprehension refers to both the activity in which the subject is involved while processing information and to the product of this activity. The product may be analyzed as a mental structure or studies through the subject’s behavior observed in different tasks. (p. 157)

What, then, is the current state of knowledge regarding these two aspects of comprehension? Of the first, Garrod (1986, p. 226) states, “psychologists are a long way from producing any definitive statement of what occurs, in terms of processing, when we actually understand an utterance in natural discourse,” and, of the second, Foss (1988, p. 303), interpreting “product” in the sense of mental structure, observes that, “We know the input to the comprehension mechanism but their product is mysterious.” Such opinions appear to be representative of the area: neither the process nor the product is adequately understood, nor is there a generally accepted theory of comprehension in cognitive psychology.

Nor, it seems, is it widely expected that such a theory will emerge in the future. This is certainly the view of Kintsch and Kintsch (1984, p. 175), who consider, quite simply, “there can be no overall theory of comprehension,” and see their own work in terms

of understanding “the specific demands and operations involved in a particular cognitive task.” They conclude their review of the area with their opinion that:

the pursuit of a new, general theory of learning... is probably just a illusionary now as it was 50 years ago. Just as there can be no general, overall theory of comprehension...we are unlikely to be able to specify a general learning process. (p. 175)

SLA researchers cannot, therefore, look to L1 cognitive psychology for an overall theoretical framework of comprehension processes.

Specifically in terms of language comprehension, the situation is much the same: the general position is described by Harris and Coltheart (1986, p. 204), who note, “there is currently no overall theory of language comprehension.”

#### **4. L2 Language Comprehension: Theory and Research**

L2 models of listening comprehension have proved no more adequate than those in L1. Nagle and Sander’s (1986) attempt to produce “a theoretical model of listening comprehension” (p. 9), for example, merely results in meaningless baptism of processes inadequately understood.

The inadequacies of theory in this area are also long-standing and continuing. Carroll (1972, p. 2), for example, observed that the 1956-59 Educational Testing Service committee were not aided in their task of publishing “STEP Tests of Listening” by the lack of a theory of listening comprehension, which might have been expected to guide the program. More recently, Brown (1986, p. 286) has commented on the inadequacies of current formulations of L2 listening comprehension in terms of practical application: “We are still a long way from developing the sort of theory of comprehension processes which is going to be much help to teachers in the classroom.”

In short, no satisfactory theory of L2 listening comprehension exists at the present time. Current attempts at theoretical formulations do, however, tend to be heavily cognitive, and discussions as to the respective contributions of top-down and bottom-up processing appear to occupy the key area of disagreement (Buck, 1988, p. 19). Discussion of the relationship between these two hypothesized

skills centers upon whether language processing is only bottom-up or whether it is both top-down and bottom-up. For some the answer appears not to be in doubt. Certainly Chomsky (in his interview with Baars, in Baars, 1986, pp. 348-349) is in no doubt: "It is pretty obvious that when you and I understand each other we're bringing in all sorts of information that has nothing to do with language. I don't think we have to show that" (p. 348).

Chomsky's dismissal of the central issue in L2 attempts at comprehension theory is perhaps overly robust and unlikely to be widely accepted, but it indicates the inherent weakness of a model in which questions are not formulated in forms which can be answered empirically and are, consequently, subject to endless debate. It is argument over this sort of model which prompted Neisser (1976, p. 8) to criticize psychologists for "lavishing too much effort on hypothetical models of the mind and not enough on analyzing the environment."

Research directed towards application in our discipline has no alternative to analyzing the environment. Chaudron and Richards (1986, p. 122), for example, describe research on listening comprehension, in both L1 and L2, as "in its infancy." This certainly describes the state of research on the role of temporal variables in L2 listening comprehension: we are not in possession of even the most fundamental facts—even, for example, the basic relationship between speech rate and comprehension by L2 learners. In L2 pausology, the state of the art, has, in fact, until recently, involved very little science (however, broadly defined). L1 has fared rather better.

### 5. Pausology: Introductory Remarks

The temporal variables most often studied in psycholinguistics are speech rate (SR) and pause phenomena (PP). The latter comprise pause duration, distribution, and frequency. In addition, the distinction between silent pauses and filled pauses is consistently employed in the literature. Hesitation phenomena (HP) have increasingly been classified, "rightly or wrongly" (Grosjean, 1980, p. 39), as temporal variables and have often been included in pausological studies. They comprise not only filled pauses (such as *er*, *um*, and the schwa) but also repeats and false starts.

Investigation of such phenomena constitutes the empirical discipline of pausology, defined by O'Connell and Kowal (1980) as, "the behavioral investigation of temporal dimensions of human speech" (p. 8). Since the mid-fifties temporal variables in speech have been intensively investigated and a body of extremely specialized research findings has resulted. Goldman-Eisler is considered to be the pioneer of pausology, and the 1978 Kassel Workshop, designated in her honor, clarified the status of pausology as a component part of psycholinguistics (see Dechert & Raupach, 1980).

Goldman-Eisler's first pausological experiments were published in the early 1950s. Since then conventions of pausological measurement have gradually been established, and the importance of comparable methodologies has been emphasized (e.g., see Kowal et al., 1983, p. 390).

Pausological investigations have, however, been largely in L1 where researchers have typically investigated speech encoding. The role of temporal variables in speech decoding has, consequently, been relatively neglected. Nor has the general standard of research on decoding been comparable to the specialized encoding studies.

Decoding investigations have been carried out in a variety of contexts: in instructional systems technology the research focus has been on the deterioration in comprehension observed to occur at rapid SRs; in "baby talk" studies, SR in caretaker-infant interaction has been observed; while, in SLA, variations in rate in "foreigner talk" have been investigated.

## **6. The Relevance of Pausological Research to Language Comprehension**

The frequency with which temporal variables are referred to in the SLA literature indicates a recognition of their importance (e.g., Hatch, 1983, p. 183; Klein, 1986, p. 45; Chaudron, 1988, p. 64). However, until recently research in this area had, almost without exception, been conducted as part of studies with wider concerns.

Yet the questions posed in pausological research are extremely relevant to language teaching and particularly important to listening comprehension. The speed of delivery, for example, when

addressing NNSs, is something that thousands of EFL teachers have to make decisions on during every lesson they teach, and HP in teacher-talk are a possible source of considerable incomprehension. These questions, and others concerning PP, represent ubiquitous issues: it cannot be wise to treat them as peripheral.

However, there has been, until recently, virtually no research basis on which, for example, recommendations of classroom SR could be made by teacher trainers. Nor has there been an adequate descriptive database from which it can be seen that modifications of SR and PP are actually used in teacher talk.

### 7. Early L2 Pausological Research

Despite Sabin et al.'s (1979, p. 54) assertion that "Efforts must... be begun on applications of pausological findings to language education," for many years (from approximately 1973-1989) the field of temporal variables has been treated as peripheral. Before 1989, in not a single article in the L2 literature had SR, PP, or HP been the sole focus of investigation, and only in one experimental study had SR been the prime focus of investigation.

A synopsis of this early work has been given by Chaudron (1985, 1988), who, despite noting that there has been very little consistent quantification of SR in the classroom (1985, p. 218), still draws a conclusion from the evidence: SR to beginner learners is about 100 wpm and SR to intermediate and advanced learners and NSs is about 30-40 wpm faster (1988, p. 69). But this is to give unmerited credence to extremely dubious data.

The studies of Henzl (1975, 1979), for example, are deficient in data collection, analysis, and interpretation. Hakansson's (1986) language classroom study is equally flawed, as is almost every study reporting on temporal variables before 1989. Moreover, differing instrumentation, methodologies, text genre, and languages mean that the studies are neither replicable nor comparable. (For detailed review see Griffith, forthcoming.)

Recently, however, published studies have begun to appear in the applied linguistics literature in which SR or PP are the major focus of investigations (Anderson-Hsieh & Koehler, 1989; Conrad, 1989; Tauroza & Allisson, in press). There has also been recent experimental work directly on SR and listening comprehension

(Griffiths, in press) and an impressive early L2 study (Grosjean, 1972) has come to light. I will begin by describing the recent pausological work, much of which has included research relevant to comprehension.

### 8. Recent L2 Pausological Research

A recently published study by Conrad (1989) focused largely on pausology: it should be seen as the first serious study of temporal variables in the SLA literature. Conrad used time-compressed speech in an experiment attempting to observe differences in aural processing strategies between NSs and NNSs of high and intermediate language proficiency. Twenty-nine NSs, 17 high-level NNSs and 11 medium-level NNSs were asked to immediately recall 5 time-compressed recordings of 16 English sentences. There were five trials for each sentence with each one being heard at 450, 320, 253, 216, and 196 wpm after which subjects were asked "to report everything and anything they thought they had heard" (p. 7).

Results showed that, whereas NSs had nearly full reports of all the sentence components by the second trial (320 wpm), NNSs of both "high" and "medium" levels of proficiency experienced considerable difficulty even with the slowest rate after the fifth hearing (at this stage the high-level group still failed to report accurately 28% of the sentence items, and the medium-level group failed to report 56%). These latter figures confirm the belief that, as with NSs, there is a level of SR at which NNSs of differing levels will experience rapid declines in comprehension. It appears from these figures that for high-level groups this may be around 200 wpm but for medium-level groups considerably less.

Another recent study (Anderson-Hsieh & Koehler, 1989) also describes the SR-comprehension relationship, but does so in relation to NSs listening to NNSs. In their investigation Anderson-Hsieh and Koehler hypothesize that, because SR studies of native speech have shown an increase in SR is associated with a decrease in comprehension, it is reasonable to expect the same relationship to be observed with nonnative speech. It is consequently investigated, as is the effect of foreign accent on comprehension.

The speakers for the study were three native speakers of Chinese of different levels of proficiency in English who between

them read six passages (310 to 475 syllables in length) reporting arcane information. A NS then read all six passages at almost exactly the same rates as the NNSs. The rates chosen were determined "empirically," that is, the slow rate was based on what the NS could read without sounding abnormally slow, and the fast rate was based on what the least proficient NNS was able to produce without sounding too rushed. This gave slow rates between 2.39 and 2.65 sps; regular rates between 3.25 and 3.49 sps; and fast rates from 4.22 to 4.58 sps. Comprehension was tested with six multiple-choice questions for each passage, various combinations of which were given to 224 NS American university students.

Results showed that scores on the passage delivered at the fast rates were significantly lower than for the regular rate for all speakers. All series of scores indicated least comprehension at the fast rate, and all but one showed greatest comprehension at the slowest rate. Significant differences in all cases were found between the fast and regular rates, and fast and slow rates, but not between the slow and regular rates.

In addition, findings from an investigation of perception of SR within the experiment (using a 5 point scale) indicated that Ss were aware of the rate differences which resulted in comprehension variability (there are somewhat mixed findings on this issue as will be indicated in the next section). Anderson-Hsieh and Koehler conclude, "the study has shown rather dramatically that speaking rate is an important factor in the comprehension of the nonnative speech investigated in this study" (p. 591). Their finding, therefore, mirrors that of Conrad's (1989) investigation which also showed L2 learners to have difficulty in understanding rapid NS speech.

Another recent study, but one which has focused on the occurrence of rate variability in different speech genres, is that of Tauroza and Allison (in press). However, their findings on SR norms could have been derived from the infinitely fuller information in the pausological literature and their recommendation that spm rather than wpm should be adopted was made by Goldman-Eisler in 1954 (p. 94). Nonetheless, the study is sounder than many previous ones in the area, and it adds to the accumulating body of knowledge on temporal variables which is now being assembled in L2 research and is now appearing in the SLA/applied linguistics

literature. It can justifiably be regarded amongst those recent contributions (in addition to Grosjean's study) which give the area a quality and an optimism which was previously lacking.

### **9. L2 Pausological Experimental Studies: Temporal Variables and Listening Comprehension**

That optimism is not, however, derived from the first study described in this section (Kelch, 1985). To date this is still the only published study on the relationship between comprehension and SR in the SLA literature. However, it is flawed in so many respects that no convincing conclusions can be derived from it. A major, indeed a fatal, flaw in the experiment was introduced when Kelch inserted extended pauses into the experimental text (originally recorded at SRs of 191 wpm and 124 wpm) in order to accommodate his chosen testing method—dictation. If these 45 sec pauses are included in the SR calculations, the delivery rate of the passages falls rather dramatically to 4.2 wpm and 4.3 wpm respectively.

A further difficulty in interpreting the results arises due to Kelch employing two very different methods of scoring. He found that while the slower SR resulted in significantly greater comprehension when measured by an exact word method, it did not do so when measured by an equivalent meaning method: in fact, the reverse relationship was observed when comprehension was measured on the different methods, but this point is not so much as mentioned in the discussion of findings. Kelch's summary of his results is, therefore, both partial and misleading and the study fails to illuminate the SR/comprehension relationship.

However, a study which is never reported in the SLA literature, but which convincingly demonstrates that modification of temporal variables can facilitate L2 comprehension, is Grosjean's (1972) unpublished doctoral thesis, briefly reported in Lane, Grosjean, Le Bette and Lewin (1973). The paper appeared in *Linguistics* and was delivered at the 1971 International Congress of Applied Psychology, "Achievements and Prospects in the Applications of Psycholinguistics to the Teaching of Foreign Languages," in Liege, but since then seems to have got lost. As just noted, however, it is a study which is well worth reviving as it is still probably the most thorough investigation in the area to date.

The study was undertaken to “see whether...differences between...high and low values of the temporal properties made a difference to comprehension—and how much of a difference” (Lane et al., 1973, p. 17). Nineteen matched pairs of adult ESL students drawn from a population of 485, heard “high” and “low” versions of 6 texts modified in terms of articulation rate (SR minus pause time), number of pauses/length of run, and pause duration. “High” and “low” in this context were defined as about 1.5 standard deviations above and below mean rates; they were, therefore, only moderately high and low. For example, on the first text “high” articulation rate was 5.46 syllables per second (sps) and the “low” rate was 4.01 sps (Grosjean, 1972, p. 57).

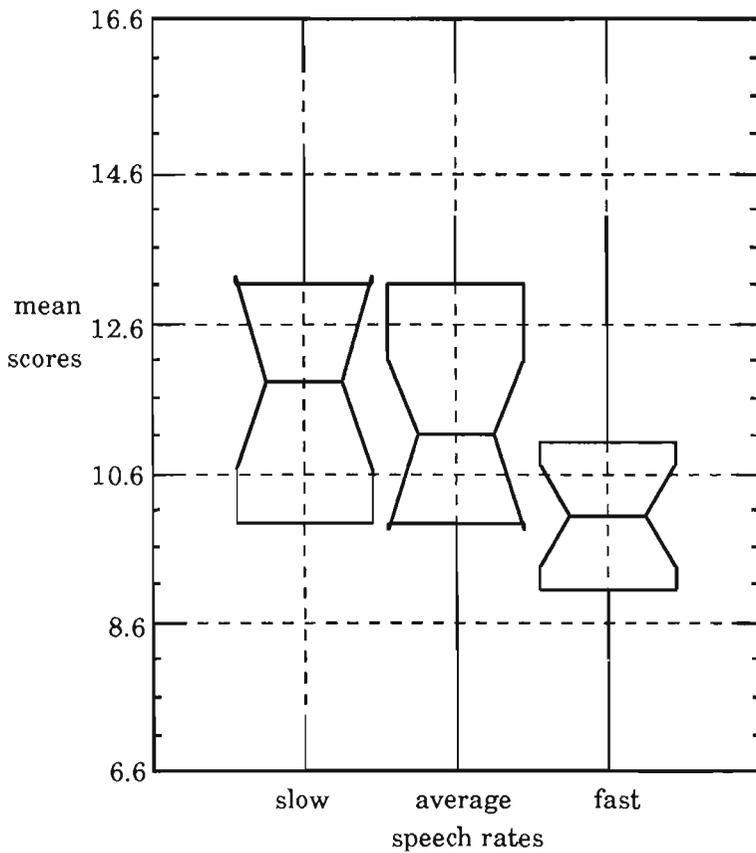
Comprehension was tested with six open-ended questions on each text. Lane et al. (1973) summarize the results in the following way: “In 5 of the 6 cases, changing from moderately high to moderately low values of a temporal variable led to moderate improvements in comprehension by non-native speakers” (p. 18). They conclude their comments on this experiment with an interesting summation of the three sets of findings:

If we simply add the effects on comprehension of the three temporal variables under study, comprehension stepped up an average of 24 per cent, but this figure depends too much on the passages used and does not take account of how the temporal variables covary normally and interact in their effects on comprehension; it is probably an underestimate of the impact of moderate changes in temporal properties on second-language comprehension. (p. 18)

Grosjean (personal communication, 6/27/89) acknowledges that summing the three sets of results may not be justified and that a different type of experimental design would be needed to investigate the combined effect of the three variables. However, even the individual results indicate the very important role played by temporal variables in facilitating comprehension. Moreover, were Lane et al.’s conclusion to be verified through replication, the implications for language teaching would be enormous: nowhere else has such a massive increase in comprehension been observed through moderately adjusting performance dimensions.

A recent investigation (Griffiths, in press) was undertaken to examine the effects of 3 SRs (200 wpm/3.8 sps; 150 wpm/2.85 sps; 100 wpm/1.9 sps) on the comprehension of three 350 to 400-word lexically and grammatically graded passages delivered to a group of 15 low-intermediate level NNSs. Repetition of the testing procedure with three texts and the use of a randomised complete block design were used to compensate for the small n and the lack of

**Figure 1**  
 Notched Box-and-Whisker Plot of Mean Scores  
 Obtained at Slow, Average and Fast Speech Rates



homogeneity amongst the subjects. The use of three texts also increased the statistical power of the experiment to acceptable levels. After several testing methods had been piloted, true-false questions were adopted to check comprehension. Two-way ANOVA indicated the observed difference in scores under the three SRs to be at the .0549 level, that is, narrowly missing the 5% level of significance. However, as this was a preliminary study conducted with a small *n*, further exploratory data analysis was undertaken (following Tukey, 1977). The notched box-and-whisker plot in Figure 1 shows not only how close the experiment came to giving a significant finding but, more importantly, shows where the difference lay within the three means.

As significant differences are indicated where notches do not overlap, it can be seen that scores obtained at the moderately fast rate and the slow rate are close to being significantly different, while the notches on the average rate box overlap both of the others. In summary, results showed that moderately fast speech rates resulted in a large reduction in comprehension, but that scores on passages delivered at slow rates did not differ markedly from those delivered at average rates.

It appears, therefore, that (in this particular context) speaking slowly did not greatly aid comprehension, but speaking moderately quickly reduced it. Should this tentative finding be replicated in future studies in different conditions, then the oft-repeated teacher-training direction to speak at normal rates (Hatch, 1983, p. 159) would, using simplified input, be empirically supported with low-intermediate students (but probably not for beginners). However, an investigation of subjective responses by both the NNS subjects and 14 NSs who also heard the recordings, showed frequent misperception of SR variation: if that is the case, consistent modification may not easily be achieved by teachers.

It should be stressed, however, that the complexity of the findings in this study and in those described above, as well as the caveats that have to accompany any conclusions that can be drawn from them, indicate the need for local studies and local applications rather than the ability of research to furnish wide-ranging recommendations with total confidence. Moreover, the value of this type of experimental research is only convincingly demonstrated when

the results are allied to those of descriptive studies (in this case of SR) conducted in the same setting.

For example, one such study (Griffiths, 1990), of the SR of 10 EFL teachers to university students of three different levels, revealed not only no significant difference between rates observed in a NS-NS baseline and rates to low-proficiency NNS learners, but also teachers talking to these students at rates of up to 4.5 sps. Even taking into account the standard caveats, it is highly unlikely that deliveries at such rates are adequately comprehended.

A related finding, from an investigation into the effect and occurrence in teacher-talk of hesitation phenomena (Griffiths, in preparation), does, however, reveal the apparently unconscious sensitivity of the above group of EFL teachers to, in particular, filled pauses. Firstly, the experimental part of the study confirmed, with low- and intermediate-proficiency NNSs, Voss's (1979) finding with advanced-level NNSs that hesitation phenomena were sources of perceptual error. A study of the occurrence of filled pauses (which caused most of the observed errors) in 30 EFL lessons showed them to be significantly less frequent ( $t[9] = 5.446$ ,  $p < .005$ ) than under comparable NS-NS baseline conditions. As it is extremely unlikely that any of the teachers in the experiment have ever given any conscious thought to the number of filled pauses in their speech to NSs or NNSs, this finding appears to show an unconscious modification of a particular temporal variable which is likely to facilitate comprehension of the input.

## 10. Summary and Conclusion

Each major category of temporal variable appears, therefore, to be capable of manipulation to facilitate NNS comprehension. If one of the major tasks of those involved in L2 teaching is simply to ensure comprehension, then these variables can be seen as amenable to manipulation. It now remains for researchers to specify more exactly how comprehension might be facilitated by manipulation of temporal variables, and to gauge how far these modifications are currently deployed in language classrooms.

However, a beginning has been made. Prior to 1989, MacLaughlin's conclusion that "We remain in a night in which all cows are black" (1980, p. 298) was the only one which could be drawn. Recent

developments require that that conclusion be changed. A number of recent investigations have not only been specifically of temporal variables but they have, to varying degrees, demonstrated the researchers' familiarity with pausological methodology. These studies have, consequently, produced more substantial findings. The study of temporal variables in an L2 context appears to be being granted the sort of prominence that it merits, and current research is doing justice to the importance of the problems being investigated.

To conclude, there is still no substantial theoretical underpinning to the research taking place and it seems unlikely that there is going to be one: as Cronbach (1982) notes generally for the social sciences: "Waiting for Newton is as pointless as waiting for Godot" (p. 61). However the detailed work which is taking place is beginning to "pin down the contemporary facts" (Cronbach, 1975, p. 126); they might then be used to inform practice. That is no mean aspiration.

*This is a revised version of a paper presented at the JALT 15th Annual International Conference on Language Teaching/Learning, Okayama, Japan, November 4, 1989.*

**Roger Griffiths** is Associate Professor of English at Nagoya University of Commerce and Business Administration. He is also a Chartered Psychologist and Associate Fellow of The British Psychological Society. He has a Ph.D. from Southampton University on the role of temporal variables in L2 learning.

### References

- Anderson-Hsieh, J., & Koehler, K. (1989). The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning*, 38, 561-613.
- Baars, B. J. (1986). *The cognitive revolution in psychology*. New York: The Guilford Press.
- Beretta, A. (1986). Program-fair language teaching evaluation. *TESOL Quarterly*, 20, 431-444.
- Brown, G. (1986). Investigating listening comprehension in context. *Applied Linguistics*, 1, 284-301.

- Brown, G. (1989). Sitting on a rocket: An interview with Professor Gillian Brown. *ELT Journal*, 43, 167-172.
- Buck, G. (1988). Testing listening comprehension in Japanese University Entrance Examinations. *JALT Journal*, 10, 15-42.
- Carroll, J. B. (1966). The contributions of psychological theory and educational research to the teaching of foreign languages. In A. Valdman (Ed.), *Trends in language teaching* (pp. 93-106). New York: McGraw Hill.
- Carroll, J. B. (1972). Defining language comprehension: Some speculations. In J. B. Carroll & R. O. Freedle (Eds.), *Language comprehension and the acquisition of knowledge* (pp. 1-29). New York: John Wiley.
- Chaudron, C. (1985). Comprehension, comprehensibility, and learning in the second language classroom. *Studies in Second Language Acquisition*, 7, 216-232.
- Chaudron, C. (1988). *Second language classrooms: Research on teaching and learning*. Cambridge: Cambridge University Press.
- Chaudron, C., & Richards J. C. (1986). The effect of discourse markers on the comprehension of lectures. *Applied Linguistics*, 7, 113-127.
- Conrad, L. (1989). The effects of time-compressed speech on native and EFL listening comprehension. *Studies in Second Language Acquisition*, 11, 1-16.
- Corder, S. P. (1984). Review of S. D. Krashen "Second Language Acquisition and Second Language Learning" and "Principles and Practice in Second Language Acquisition." *Applied Linguistics*, 5, 56-58.
- Cronbach L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, 30, 116-127.
- Cronbach, L. J. (1982). Prudent aspirations for social inquiry. In W. M. Cruskal (Ed.), *The social sciences: Their nature and uses* (pp. 61-81). Chicago: Chicago University Press.
- Cronbach L. J. (1987). Social inquiry by and for earthlings. In *Evaluation Studies* (Vol. 12, pp. 400-424). Newbury Park: Sage.
- Dechert, D. W., & Raupach, M. (Eds.). (1980). *Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler*. The Hague: Mouton.
- Ehrlich, M-F. (1982). An experimental study of the relationship between comprehension and memorization of a text. In J-F Le Ny & W. Kintsch (Eds.), *Language and comprehension* (pp. 157-168). Amsterdam: North Holland.
- Ellis, R. (1985). *Understanding second language acquisition*. Oxford: Oxford University Press.
- Eysenck, H. J. (1986). Consensus and controversy—Two types of science. In S. Modgil & C. Modgil (Eds.), *Hans Eysenck: Consensus and controversy* (pp. 375-398). Philadelphia, PA: The Falmer Press.
- Feyerabend, P. (1970). Consolations for the specialist. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 197-230). Cambridge: Cambridge University Press.
- Foss, D. J. (1988). Experimental psycholinguistics. *Annual Review of Psychology*, 39, 301-348.

- Garrod, S. (1986). Language comprehension in context: A psychological perspective. *Applied Linguistics*, 3, 226-238.
- Goldman-Eisler, F. (1954). On the variability of the speed of talking and on its relation to the length of utterances in conversation. *British Journal of Psychology*, 45, 94-107.
- Gregg, K. R. (1984). Krashen's Monitor and Occam's Razor. *Applied Linguistics*, 5, 79-100.
- Gregg, K. R. (1986) Review of "The input hypothesis: Issues and implications" by S. D. Krashen. *TESOL Quarterly*, 20, 116-122.
- Gregg, K.R. (1988). Epistemology without knowledge: Schwartz on Chomsky, Fodor and Krashen. *Second Language Research*, 4, 66-80.
- Griffiths, R. T. (1990, March). *Language classroom speech rates: Life in the fast lane*. Paper delivered at the 24th Annual TESOL Conference, San Francisco, CA.
- Griffiths, R. T. (in press). Speech rate and NNS comprehension: A preliminary study in time-benefit analysis. *Language Learning*, .
- Griffiths, R. T. (forthcoming). Pausology and SLA. *Applied Linguistics*.
- Griffiths, R. T. (in preparation). The paradox of comprehensible input: Hesitation phenomena in L2 teacher-talk.
- Grosjean, F. (1972). *Le rôle joué par trois variables temporelles dans la compréhension orale de l'anglais étudié comme seconde langue et perception de la vitesse de lecteurs et des auditeurs*. Doctoral dissertation. University of Paris V11.
- Grosjean, F. (1980). Temporal variables within and between languages. In H. W. Dechert & M. Raupach (Eds.), *Towards a cross linguistic assessment of speech production* (pp. 39-53). The Hague: Mouton.
- Håkansson, G. (1986). Quantitative aspects of teacher talk. In G. Kasper (Ed.), *Learning, teaching and communicating in the foreign language classroom* (pp. 83-98). Aarhus: Aarhus University Press.
- Harris, M., & Coltheart, M. (1986). *Language processing in children and adults: An introduction*. London: Routledge & Kegan Paul.
- Hatch, E. M. (1983). *Psycholinguistics: A second language perspective*. Rowley, MA: Newbury House.
- Henzl, V. M. (1975, August). *Speech of foreign language teachers: A sociolinguistic register*. Paper delivered at the 4th International Congress of AILA, Stuttgart.
- Henzl, V. M. (1979). Foreign talk in the classroom. *IRAL*, 17, 159-167.
- Kasper, G. (1988). Editor's Note. In G. Kasper (Ed.), *Classroom Research—AILA Review*, 5 (pp. 5-9). Amsterdam: Free University Press.
- Kelch, K. (1985). Modified input as an aid to comprehension. *Studies in Second Language Acquisition*, 7, 81-90.
- Kintsch, W., & Kintsch, E. (1984). Studies in test comprehension: Toward a model for learning from reading. In V. Sarris & A. Pardicci (Eds.), *Perspectives in psychological experimentation: Toward the year 2000* (pp. 165-176). Hillsdale, NJ: Lawrence Erlbaum .
- Klein, W. (1986). *Second language acquisition*. Cambridge: Cambridge University Press.
- Kowal, S. H., Wiese, R., & O'Connell, D. C. (1983). The use of time in storytelling.

- Language and Speech*, 26, 377-392.
- Krige, J. (1980). *Science, revolution and discontinuity*. Brighton: Harvester.
- Kuhn, T. S. (1970a). *The structure of scientific revolutions*. Second Edition. Chicago: University of Chicago Press.
- Kuhn, T. S. (1970b). Reflections on my critics. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 231-278). Cambridge: Cambridge University Press.
- Lane, H., Grosjean, F., Le Berre, J., & Lewin, E. (1973). Exploring some properties of foreign-language utterances that control their comprehension. *Linguistics*, 112, 15-22.
- Lightbown, P. M. (1985). Great expectations: Second language acquisition research and classroom teaching. *Applied Linguistics*, 6, 173-189.
- Long, M. H. (1985a). Input and second language acquisition theory. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 377-393). Rowley, MA: Newbury House.
- Long, M. H. (1985b, February). *Theory construction in second language acquisition*. Paper presented at the Second Language Research Forum, University of California, Los Angeles.
- Long, M. H., Brock, C., Crookes, G., Deicke, C., Potter, L., & Zhang, S. (1984). *The effect of teachers' questioning patterns and wait-time on pupil participation in public high school classes in Hawaii for students of limited proficiency*. Manoa, Hawaii: University of Hawaii at Manoa.
- McLaughlin, B. (1978). The monitor model: Some methodological considerations. *Language Learning*, 28, 309-332.
- McLaughlin, B. (1980). Towards a theory of speech processing: Some methodological problems. In D. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech: Studies in honour of Freida Goldman-Eisler* (pp. 291-298). The Hague: Mouton.
- McLaughlin, B. (1987). *Theories of second language learning*. London: Edward Arnold.
- Masterman, M. (1970). The nature of a paradigm. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 59-89). Cambridge: Cambridge University Press.
- Medawar, P. (1984). *The limits of science*. Oxford: Oxford University Press.
- Nagle, S. J., & Sanders, S. L. (1986). Comprehension theory and second language pedagogy. *TESOL Quarterly*, 20, 9-26.
- Neisser, U. (1976). *Cognition and reality*. San Francisco: Freeman.
- O'Connell, D. C., & Kowal, S. (1980). Prospectus for a science of pausology. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler* (pp. 3-10). The Hague: Mouton.
- O'Hear, A. (1989). *Introduction to the philosophy of science*. Oxford: Clarendon Press.
- Popper, K. R. (1970). Normal science and its dangers. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 51-58). Cambridge: Cambridge University Press.
- Sabin, E. J., Clemmer, E. J., O'Connell, D. C., & Kowal, S. (1979). A pausological

- approach to speech development. In A. W. Siegman & S. Feldstein (Eds.), *Of speech and time* (pp. 35-55). Hillsdale, NJ: Lawrence Erlbaum.
- Sinclair, J. M. (1989, November). *The divorce of practice and theory*. Paper delivered at the 15th JALT Conference, Okayama, Japan.
- Tauroza, S., & Allison, D. (in press). Speech rates in British English. *Applied Linguistics*.
- Tukey, J. W. (1977). *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.
- Voss, B. (1979). Hesitation phenomena as sources of perceptual errors for non-native speakers. *Language and Speech*, 22, 129-144.

## Where Do Tests Fit Into Language Programs?

**James Dean Brown**

*University of Hawaii at Manoa*

The central thesis of this paper is that testing should be used to guide the development of a sound language program. The paper begins by examining the four types of decision making processes used in any language teaching institution: (a) who should be admitted into the program; (b) which level is appropriate for each student; (c) what should be taught so that resources are maximally utilized; (d) which students should be promoted to the next level? In order to help staff make these decisions, four categories of tests (one for each type of decision) are discussed: proficiency, placement, diagnostic, and achievement. Each is examined in terms of the purpose of the decision (norm-referenced or criterion-referenced), and the type of information that it provides. The paper concludes with a discussion of a hypothetical case study. The main point is that testing is not the enemy. On the contrary, tests can provide guidance in making informed and responsible decisions.

### 言語教育におけるテストの位置付け

本研究は、主に、言語プログラムの発展・充実に際して、テストが活用されるべきである、ということについて論ずるものである。まず、いかなる教育機関においても決定を下す過程で採用し得る、以下に掲げる4つの事項について検討をする。

1. そのプログラムに入れるのは誰であるべきなのか。
2. 個々の学生に適したレベルはどれなのか。
3. 学生の力を最大限に伸ばすためには何を教えるべきなのか。
4. 上のレベルに進めるのはどの学生であるべきなのか。

これらの決定を下す際の参考に供するために、実力 (proficiency)、クラス分け (placement)、診断 (diagnostic)、成就 (achievement) の4種類のテスト (1つの事項に1つのテスト) を取り上げる。さらに、そのテストの提供し得る情報の種類及び、(全般的な能力もしくは、学習範囲内での能力といったことも考慮して) 決定を下す際の妥当性といった観点からそれぞれのテストに検討を加える。締めくくりとして、ある学習者を想定し、そのケースに沿って、いかに本論が適用し得るかを論ずるが、肝心なのはテストは「敵」なのではなく、情報に富み、妥当な判

断を下す上での有効なガイドラインになり得るということである。

## 1. Introduction

Many language teachers dislike the whole idea of testing. This distaste may be due to the fact that testing is so often viewed as a “necessary evil” rather than something that we do because we want our students to learn. This paper tries to turn that idea around. The central theme is that testing can and should be used to guide the development of a sound language program and help the students, as well. Thus the paper begins by examining some of the very real-life decision making processes that go on in any language program, such as deciding:

1. who should be admitted into the program;
2. which ability grouping, or level, is appropriate for each student;
3. what should be taught so that resources are most effectively allocated;
4. and, which students have accomplished enough to be advanced to the next level of study.

Such decisions are of particular importance in students' lives and should therefore be based on the best available information. Language tests are an important source of that information.

Hence, another main topic of discussion will be the types of tests that are best suited for helping teachers and administrators make decisions. Four categories of tests will be defined and discussed here. They are (a) proficiency tests, (b) placement tests, (c) diagnostic tests, and (d) achievement tests. These decision making procedures will be examined in terms of the purpose of each type of decision and the type of information that is needed. In addition, examples will be drawn from my personal experience to suggest a number of ways that the different types of tests can be effectively implemented, sequenced, and administered. As always, it will be necessary to consider some practical and political realities as well—realities that exist in any language program, particularly with regard to its tests.

The main point here is that testing is not just a “necessary evil.”

On the contrary, tests, if effectively used, can assist language teachers and administrators in making informed and responsible decisions about students, curriculum, and future policy.

## 2. Two Families of Decision Making

Let us start by thinking about the processes involved in decision making. Administrators may feel that there are infinite decisions that must be made on a daily basis. These vary from minor decisions like what kind of paper clips to buy, to rather major ones like what theoretical directions a program should take. On the other hand, teachers often feel that the really important decisions are those directly related to teaching and to students' learning processes. When I am teaching, decisions about paper clips or even about theoretical perspectives suddenly seem unimportant to me. However, there are administrative decisions that bear upon teaching and vice versa. The focus, here, will be on those types of truly important and practical decisions that affect the lives of all participants in a program: the administrators, the teachers, and the students alike.

**Figure 1**  
Decision Making Using the Four Testing Types

Level of Decision	Program Level		Classroom Level	
Type of Decision	Admissions or Institutional Comparisons	Grouping Ss into Class Levels	Grading, Promotion & Graduation	Diagnosis of Strengths & Weaknesses
Family of Test	Norm-Referenced		Criterion-Referenced	
Type of Testing	Proficiency Testing	Placement Testing	Achievement Testing	Diagnostic Testing

As shown in Figure 1, there are at least four major types of decisions that affect an entire program. Two of these are the primary responsibility of the administrators. These are essentially

questions about who should be admitted into the program and about which ability grouping, or level, is most appropriate for each student. These will be called *program level decisions* because they have to do with getting the students into the program and into appropriate courses. Such decisions must fit an unknown student into a known program. For example, if students have had previous training at other institutions, it is first necessary to decide if their level of general language ability is too low, too high, or just right for them to fit into the known program. This is the nature of admissions decisions.

Once it has been decided that certain students belong in the program, it is necessary to determine what level within the program is most appropriate for each of them, given their previous training and the training that is available in the program. Should the students be put into the elementary, intermediate, or advanced level? Should they be put into the reading course or listening, writing, and speaking courses? Such decisions are most often entrusted to the administrators who are responsible for the logistics of registration and enrollment in the program. Nevertheless, the teachers will ultimately work most closely with the students, and therefore should be involved in these types of decisions, even if the primary obligation rests with the administrators.

Other decisions (also shown in Figure 1) are the primary responsibility of the teachers. These are usually questions about what should be taught in a course, about students' progress, or about which students have accomplished enough at the end of the course to be advanced to the next level of study. These will be labeled *classroom level decisions* because they have to do with the teaching-learning processes. Such decisions are based on behavior that takes place in the classroom, where the teacher knows best what is going on. Therefore, these decisions are most often made during or at the end of instruction, and are almost entirely internal to the program. In short, they are decisions that help the students progress smoothly through the program.

For example, once the students are in a particular course, it is necessary to make judgments about who needs to work hard on which objectives. Such decisions are usually made early in the course in order to help students use their energies most efficiently. At the midterm, decisions about student progress may require

assigning a grade or otherwise reporting to the administration. At the end of a course, evaluations of each student's success in the course may be necessary, either in terms of a grade or in terms of promotion to the next level of study, or both. They are classroom level decisions and should be primarily the responsibility of the teachers. Nevertheless, the administrators should provide as much assistance to the teachers (in terms of test design and logistics) as possible—even though the primary responsibility for classroom decisions should always remain with the teachers.

In my work, I have repeatedly found that well-designed tests provide one source of valuable information for decision making. I have also found that the two families of decisions, program level and classroom level, are aided by two more-or-less parallel families of tests, called norm-referenced and criterion-referenced tests. In all cases, the tests in our English Language Institute (ELI) have been developed cooperatively with input from administrators, teachers, and students. The goal of our testing program is to help with the types of program level decisions that the director and assistant director must make, as well as the classroom decisions that the teachers must make. Let us consider each of these test types separately.

### 3. Two Families of Language Tests

Perhaps the most important thing to remember about language tests is the fact that the results can be viewed from a variety of different perspectives and, therefore, can serve a variety of decision making functions within any language program. The problem is matching the correct type of test with the type of decision. To this end, let us consider the two families of tests, *norm-referenced* and *criterion-referenced*, in much more depth.

These two terms are relatively new in our field (see Bachman, 1989; Brown, 1984a, 1988, & 1989 a & b; Cziko, 1983; Davidson, Hudson, & Lynch, 1985; and Hudson & Lynch, 1984), though advocated for years in some educational testing circles (see Berk, 1980; and Popham, 1978 & 1981, for much more on this topic). In fact, the contrast between norm-referenced and criterion-referenced tests dates back to Glaser (1963) and is the focus of many articles and studies in educational and psychological testing jour-

nals (see any recent issue of *Journal of Educational Measurement* or *Applied Psychological Measurement*). The distinction is of increasing importance in those circles, and hopefully will become more significant in language teaching. Thinking about tests in these two distinct ways can help us understand the differences and similarities between the various types of tests that are administered, and the decisions that must be made.

**Table 1**  
Differences Between Norm-referenced and Criterion-referenced Tests

Characteristic	Norm-referenced	Criterion-referenced
1. Type of interpretation	Relative: A student's performance is compared to that of all other students by that student	Absolute: A student's performance is compared only to the amount, or percent of material learned
2. Type of information	General Language abilities or proficiencies	Specific Objectives-based language points
3. Purpose of Testing	To spread students out along a continuum of general abilities or proficiencies	To assess the amount of material known, or learned by each
4. Score Distribution	Normal distribution of scores around a mean score	Ideally, if students know all of the material, they should 100%
5. Knowledge of questions	Students have little or no idea of what content to expect in questions	Students know exactly what content to expect in test questions

(adapted from Brown 1984a, 1989a)

There are a number of important contrasts between norm-referenced tests and criterion-referenced tests in terms of both the theory and practice of language testing. Some of the practical

concerns are summarized in Table 1, which shows that the two sorts of tests differ in the (a) ways scores are interpreted; (b) types of information gathered; (c) purposes for testing; (d) score distributions; and (e) knowledge that students have beforehand about the items.

In brief, norm-referenced tests (NRTs) are designed to measure global language skills or abilities (e.g., overall English language proficiency, reading comprehension, and so forth). A student's score on a norm-referenced test is interpreted in relation to the scores of the other students who took the test. This interpretation is typically done using the statistical concept of normal distribution (the "bell curve") of scores dispersed around the mean, or average, sometimes in terms of percentiles or other standardized scores. The purpose of a norm-referenced test is to spread students out along a continuum of scores so that those with "low" abilities end up at one end of the normal distribution, while those with "high" abilities are found at the other (with most of the students falling somewhere between the extremes near the mean). On a norm-referenced test it is also common that, even though the students may know the general form that the examination questions will take (i.e., the types of items), they typically do not know what specific content will be tested by those questions. For example, they may know that the questions will be multiple-choice, true-false, and so forth, but they usually have no idea what the questions will actually test, except in the broadest terms.

Criterion-referenced tests (CRTs) are produced to measure clearly defined instructional objectives. Often those objectives are unique to a specific program and serve as the basis or outline of the curriculum. Thus, it is usually important for the teachers and students to know exactly what the objectives are so that suitable amounts of time can be spent on them. The interpretation of criterion-referenced test scores is absolute in the sense that each student's score is meaningful by itself, without reference to any other students' scores. In other words, a student's score on a particular objective indicates the percentage of the skill or knowledge in that objective which has been learned or acquired. Furthermore, the distribution of scores on a criterion-referenced test will not necessarily be normal. In theory, if every student has learned 100 percent of each objective, it follows that all the students will

attain the same score. Finally, the purpose of criterion-referenced tests is to measure the degree to which students have developed skills or knowledge in relation to a specific set of objectives. Hence, the students should know what to expect on the test in terms of question types, as well as in terms of the tasks and content for each objective. This information would probably be implied, if not explicitly stated, in the objectives of the course.

The discussion here of norm-referenced and criterion-referenced tests has centered on practical and important differences in the type of measurement involved, the way scores are interpreted and distributed, the purpose for giving each type of test, and the students' knowledge of question content. There are also numerous contrasts between norm-referenced tests and criterion-referenced tests in the ways that they are viewed empirically and treated statistically (see Brown, 1989c; Hudson & Lynch 1984), but for the purposes of this paper, this basic description of the practical differences between the two families of tests will suffice.

#### **4. Matching Tests to Decision Purposes**

So far, the discussion has centered on two sorts of decisions that must be made in almost any language program and has shown that there are two families of tests that can be used to help us make those decisions. Before deciding which specific type of test to use for a particular type of decision, it is essential to carefully consider the true purposes of making that decision. Only then is it possible to match the correct type of test to that purpose. In this section, suggestions will be made for ways of making such matches within the context of a language program. It will begin by reviewing the four most commonly used varieties of language tests, while summarizing the main points that should be kept in mind when matching the correct variety of test to the type of decision to be made (see Table 2).

Over the years, the four categories which have been most prominent in language testing are proficiency, placement, achievement, and diagnostic tests (e.g., Alderson, Krahnke, & Stansfield, 1987, pp. iii-iv). As shown in Table 2 these four types of language tests can be categorized very neatly into the two families of tests just discussed: norm-referenced tests tend to be more useful in

**Table 2**  
Matching Tests to Decision Purposes

Test Qualities	Type of Decision			
	Norm-referenced		Criterion-referenced	
	Proficiency	Placement	Achievement	Diagnostic
Detail of Information	Very general	General	Specific	Very specific
Focus	General skills pre-requisite to entry	Abilities, levels and skills in particular program	Objectives of the course or program	Strengths & weaknesses in course objectives
Purpose of Decision	To compare individuals with other groups/ individuals	To find each student's appropriate level	To determine the degree of learning vis-à-vis program objectives	To inform student and teachers of objectives needing more work
Relationship to Program	Comparisons with other institutions	Comparisons within program	Directly related to achievement of program objectives	Directly related to progress on program objectives
When Administered	Before entry and sometimes at exit	Beginning of program	End of courses	Beginning and/or middle of courses
Interpretation of Scores	Percentile position of scores	Percentile position of scores	Amount or percent of objectives learned	Amount or percent of objectives known

making program level decisions (i.e., proficiency and placement); and criterion-referenced tests, are most effective in helping to make classroom level decisions (i.e., diagnostic and achievement). Let us consider each of these categories separately.

#### *4.1 Program Level Decisions: Proficiency and Placement*

*Proficiency decisions: Who should be admitted into the program?*

Sometimes in making decisions, we need to know the students' general levels of language proficiency. The focus of such decisions

is usually on the general knowledge or skills prerequisite to entry into, or exit from, some type of institution, for example, an American university. Making such proficiency decisions may be necessary in setting up entrance and exit standards for a program, in adjusting the level of program objectives to the students' abilities, or in making comparisons across programs. In other words, a variety of curricular and administrative questions may be usefully answered on the basis of overall proficiency information.

One way to gather information for these types of decisions is to compare the overall language performances of individuals to those of other individuals or groups. For this reason, proficiency decisions are often based on proficiency tests specifically designed for such decisions. These tests are constructed to assess general skills commonly required or prerequisite for entry into (or exemption from) a group of similar institutions. One example is the Test of English as a Foreign Language (TOEFL), which is used by many American universities that have English language proficiency prerequisites in common (see ETS, 1987). Such a test will necessarily be very general in nature and cannot be specific to any particular program. Another example of just how general such a test can and must be is found in the ACTFL Proficiency Guidelines (ACTFL, 1986). Though this type of test may contain subtests for each skill, the approach to these skills is still very general and the resulting scores can only be used as overall indicators of proficiency.

Since proficiency decisions are based on knowing the general level of language students in comparison to other students, a test is needed which will provide scores that form a wide distribution so that our interpretations of the differences between students can be fair. Thus proficiency decisions are best based on norm-referenced tests because norm-referenced tests have all of the qualities desirable for such decisions. Proficiency decisions may sometimes seem unfair, but they are often necessary to protect the integrity of the institution, as well as to protect the students from entering a program beyond their abilities or one which they really do not need.

Proficiency decisions include determining how well arriving students will fit into the program, how well the goals of the program meet the actual language needs of the students, and whether students are learning enough, as they go along, to meet overall program goals. When students leave, such tests can be used to

decide whether their level of language proficiency is high enough for their purposes, or alternatively, to discover if the program has had little impact on the language skills that students will be using in the real world. Either way, the information may prove useful to overall curriculum revision.

There are also times when comparisons are made across programs. Since proficiency tests, by definition, are general in nature, they can be used to compare regional branches of a particular language program or to compare different language centers nationwide. It is essential that such decisions be made with extreme care because serious problems can arise. These problems are often due to the very quality that makes such comparisons possible at all, that is, the fact that such tests are not geared to any particular language program. Depending on the test involved, this may result in one program which has students who score very low on the test because the teaching and learning that is going on (though perfectly effective and useful) is simply not assessed by the test. Hence such comparisons must be made with conscientious attention to the validity and appropriateness of the tests for the decisions being made. As with all tests, the information should be assessed carefully and used as part of a larger overall system of information gathering and decision making (see Brown, 1989b).

The general nature of proficiency decisions makes it essential that such tests be designed so that the general abilities or skills of students are reflected on a continuum, or spread of scores. This is desirable so that comparisons among students, or groups of students, can be rationally made. This need for a spread of scores most often leads to the construction of a test that produces a normal distribution (bell curve). All of this is to say that proficiency tests are, and most often should be, norm-referenced in purpose and character.

At the University of Hawaii, proficiency scores, in the form of TOEFL results, are used for two purposes: admission to the university and exemption from ELI training. Students must score a minimum of 500 on the TOEFL to be admitted to studies at our university (some individual departments require more). Once admitted, all international students are subject to clearance from the ELI. In many cases, that means that they must take the ELI Placement Test. However, students who scored 600 or higher on

the TOEFL are automatically exempt from ELI study.

Proficiency decisions should not be dismissed lightly. On the contrary, they must be based on the best obtainable proficiency tests as well as on other types of information. These are decisions that can dramatically affect the students' lives, so it would be grossly unprofessional to do a slipshod job with the proficiency aspect of any language program.

*Placement decisions: Which ability grouping, or level, is appropriate for each student?*

Still relatively general in purpose, placement decisions are those made for grouping students of similar ability levels together. Often, this is so that teachers can focus in each class on the problems and learning points appropriate to students at a more or less homogeneous level. To help in making such decisions, there is a category of tests designed to aid in deciding what each student's appropriate level will be within a specific program, skill area, or course. The purpose of such tests, then, is to discover which students have more of, or less of, a particular ability, knowledge, or language skill than other students in a particular program.

In thinking about placement tests, it is important to consider the similarities and differences between proficiency and placement testing. As they are defined here, proficiency and placement tests may look very similar and may both be general in nature. However, proficiency tests are designed to assess extremely wide bands of abilities. Hence they are usually applicable across a wide array of institutions. Placement tests, on the other hand, must be more specifically related to a given program, particularly in terms of the range of abilities assessed, so that students can be efficiently separated into approximately homogeneous groups which can effectively be taught similar levels of material.

In other words, a general proficiency test might be useful for determining which language program is most appropriate for a student. Once in the program, the student would take a placement test so that his or her level of study could be determined. Different levels might be appropriate for different skills.

Both proficiency and placement tests should be designed as norm-referenced instruments so that decisions can be made on the students' relative abilities or skill levels. However, as demon-

strated in Brown (1984b), the degree to which a test is effective in spreading students out along a continuum is directly related to the degree to which that test fits the ability levels of the students. Hence it is important to remember that a proficiency test will typically be norm-referenced to a population of students with a very wide band of language abilities and a variety of purposes for learning the language. This is the case with TOEFL, and with the ACTFL guidelines. However, a placement test must be norm-referenced to a narrower band of abilities and purposes—usually those found in students at the beginning of studies in a particular language program.

This distinction becomes particularly important in programs which have tracks and levels. For instance, in the ELI at Hawaii, students who arrive have already been fully admitted. In order to be admitted, they have taken the TOEFL and scored at least 500. In other words, it has been decided in language proficiency terms that these students are eligible to study in the ELI and simultaneously take a few courses at the University. Those students who score 600 or above on the TOEFL are informed that they are completely exempt from ELI training. Thus, with only rare exceptions, most ELI students have scored between 500 and 600 on the TOEFL.

Within the ELI, there are three tracks, each of which is focused on one skill (reading, writing, or listening). Within those skill areas, there are also levels: two levels each for reading and listening, and three levels of writing instruction. As a result, the placement decisions and the tests upon which they are based must be much more focused than the information provided by TOEFL scores. The placement tests must provide information on each of the three skills involved, as well as on the language needed by students in the relatively narrow proficiency range reflected in their TOEFL scores (see Brown, 1987a & b for more details on the entire placement process in the ELI).

There is a dramatic difference between our general proficiency decisions and our placement decisions. While the contrasts between proficiency and placement decisions may not be quite so clear in all programs, the ELI's definitions, and the way we distinguish between proficiency and placement decisions, may prove useful in thinking about testing in other institutions.

In the case of placement, it is particularly important to examine each test in terms of how well it fits the abilities of the students and how well it matches what is actually taught in the classrooms. If there is a mismatch between the tests and the curriculum (see Brown, 1981), the placement of students into levels may be based on something entirely different from what is taught in the levels of the program. So, typically, it is a good idea to make sure that placement decisions are based on placement tests which are either designed or adapted for the specific program involved, or, at least, seriously examined for their appropriateness for the particular program (see Brown, 1989a).

#### *4.2 Classroom Level Decisions: Achievement and Diagnosis*

*Achievement decisions: Which students have accomplished enough to be promoted?*

In a sense, we are all in the business of fostering achievement in the form of language learning. In fact, the purpose of most language programs is to maximize the possibilities for students to achieve a high degree of language learning. As a result, a time always arrives when teachers become interested in making achievement decisions. These are decisions related to the achievement (that is, the amount of learning) of our students. Teachers may also find themselves wanting to make rational decisions that will help improve achievement in the form of deciding or justifying changes in curriculum, staffing, facilities, materials, equipment, and so forth.

In order to make such decisions about student achievement and how to improve it, tests are essential. For instance, achievement tests may help in discovering how much of the language material in the program has been absorbed by each person. Thus achievement tests must be designed with specific reference to a particular program. This link with a specific program usually means that the achievement tests will be directly based on program objectives. Such tests will typically be given at the end of a course or program to determine how effectively students have mastered the objectives.

The tests used to monitor such achievement must be flexible in the sense that they can readily be made to change in response to what is learned from them about the other elements of the curricu-

lum. Within the ELI at the University of Hawaii, criterion-referenced achievement tests have been developed for each of our courses. In each case, the test items were developed directly from the objectives of the course involved. However, in the process of constructing and using these tests, we have discovered many things, not just about the students' learning, but also about the appropriateness of the objectives for those students and about the effectiveness of instruction in each of the objectives.

In other words, well thought out achievement decisions are ones based on tests from which a great deal can also be learned about the curriculum and the program as a whole. These tests should, in turn, be very responsive in the sense that they must be used to affect changes and continually test those changes against the program realities.

*Diagnostic decisions: What should be taught so that resources are most effectively allocated?*

From time to time, we may also be interested in assessing the strengths and weaknesses of each individual student vis-à-vis the instructional objectives for purposes of correcting the individual's weaknesses "before it is too late." Diagnostic decisions are aimed at fostering achievement by promoting the strengths and eliminating the weaknesses of individual students. Naturally, the curriculum is designed for the entire group of students, but in practice, attention is given to each individual. Thus, with this type of test, it is sound practice to report the performance level on each objective (as a percent) to individual students so that they can decide how and where to invest their time and energy most profitably.

Remember that this last category of decisions is concerned with diagnosing problems that students may be having in the learning process. While this type of decision is definitely related to achievement, diagnostic testing generally requires more detailed information about the specific areas in which students have strengths and weaknesses. The purpose is to help students and their teachers to focus their efforts effectively.

As with achievement tests, diagnostic tests are designed to determine the degree to which the specific instructional objectives have been accomplished. While achievement decisions are usually focused on the degree to which these objectives have been accomplished at the end of the program or course, diagnostic decisions are

normally made as the students are learning the language. As a result, diagnostic tests are typically administered at the beginning or in the middle of a course. In fact, if well constructed to reflect the course objectives, one criterion-referenced test in three equivalent forms could serve as a diagnostic tool at the beginning and mid-points in a course, and as an achievement test at the end.

Within the ELI at the University of Hawaii, the criterion-referenced achievement tests discussed above have been developed in two forms (A and B) for each course so that they can be administered at the beginning and end of each course. At the beginning, randomly selected students take Form A and the remaining students take Form B. At the end of the course, all students take the form they did not take earlier. This type of "counterbalanced" design insures that each student takes a relevant pre-test and post-test without taking the same one twice. It also allows us to diagnose potential areas of student difficulty by analyzing the results of the criterion-referenced tests, particularly the pre-test results.

Again, this system of a diagnostic/achievement testing is important, not only to provide information for the students and teachers, but also to improve the tests themselves, and to implement effective language curriculum planning and development (see Brown 1989b). Such tests are particularly important in examining and revising the program goals and objectives. The information gained can be useful in reexamining the language learning needs of the students, in selecting or creating materials and teaching strategies, as well as in evaluating program effectiveness. Thus it can be argued that the development of systematic diagnostic and achievement tests is crucial to the development of a systematic curriculum. A needs analysis is just a needs analysis and objectives are just so many notions unless they are implemented and tested against the realities of the language learning situation.

## 5. A Case Study

To review the testing processes described above, let us take a slightly different point of view for a moment. The ultimate user of the services delivered by any language program is the student. Let us consider for a moment a hypothetical student, Toshi, who wants

to study for a Master of Science degree in Oceanography at the University of Hawaii.

In order to be admitted to the program, Toshi must fill out an application, get letters of reference, have transcripts of previous studies sent to the university, and demonstrate adequate English language proficiency by taking the TOEFL. Once his application is complete and TOEFL results are received, the Graduate Division and Oceanography Department are responsible for determining whether or not Toshi is academically admissible—including whether he has a high enough TOEFL score. As mentioned above, the lowest TOEFL score allowed is 500, and Toshi scored 561.

Once Toshi is admitted to his degree program, the TOEFL proficiency test results are sent to the ELI. If he had scored over 600, he would automatically have been exempt from ELI training. However, since he did not, we must consider him for further training in ESL. In his acceptance letter, Toshi is told that he must report to the ELI as soon as he arrives in Honolulu. When Toshi reports to the ELI office, he discovers that he must take the English Language Institute Placement Test (ELIPT) three days later. He signs up for the test and receives a pamphlet describing the various parts of the ELIPT. Toshi reads this information so that he will know what to expect.

He arrives at the placement examination at 7:30 in the morning and takes the following six subjects: Academic Listening Test, Dictation, Reading Comprehension Test, Cloze, Writing Sample, and Academic Writing Test (see Brown, 1987b). The ELIPT is administered and scored primarily by the ELI teachers, so this is Toshi's first contact with that staff. He is finished with the test battery about 11:30 and goes for lunch. At 1:30, he returns for a placement interview with one of the teachers. At that time, Toshi is told that his scores indicate that he is exempt from the listening and reading courses, but that he must take the advanced writing course for graduate students (ELI 83). He agrees, and signs up for the ELI course at the same time that he registers for his oceanography courses.

During the second week of his ELI class, he takes another test. Toshi notices that it is Form B. On the basis of this diagnostic pretest, Toshi's particular strengths and weaknesses can be determined. Perhaps he is told that his grammar is very good, but that

he should concentrate on improving his organization and mechanics. Toshi studies advanced writing for 15 weeks. In the process, he is required to learn word processing and gets a great deal of practice in writing, proofreading, and revising in English.

At the end of the course, he must take another test, this time for achievement. Toshi notices that the test is very similar to the one he took at the beginning of the course, but different too. He also notices that it is labeled Form A. He does very well on this achievement test; that is good because it counts for 20 percent of his grade. Toshi's overall performance in the course, including achievement, is recorded by the teacher and reported to his Oceanography Department advisor. Toshi has completed all ELI requirements.

However, two years later, when he is finishing his master's degree, he decides to apply for a Ph.D. program at Scripts Institute near San Diego, California. It also requires information about his overall English language proficiency in the form of a TOEFL score. Because Toshi has taken the ELI course and spent two years in the United States, his score on the TOEFL is much better. He gets a score of 617, which tells the ESL professionals associated with Scripts that Toshi should be exempt from any further ESL training.

The central message in this hypothetical case study is that the student's viewpoint is important. We must always remember that it is the student who will be most affected by our tests and decisions. It is also important to note that, though there was a great deal of testing involved, it all seemed quite natural. That will be true if the testing is set up as an integral part of the overall curriculum.

## 6. Conclusion

Clearly, test results can be viewed from a variety of perspectives and serve a variety of different decision making functions within a language program. All language tests are actually based on one of two completely different families of tests: norm-referenced tests designed to help make program level determinations (like proficiency and placement decisions), and criterion-references tests constructed to help make classroom level judgments (like diagnostic and achievement decisions). There are also four primary decision making functions that tests can serve (proficiency, placement, diagnostic, and achievement). We have explored how we might best go about matching our tests to the purposes and decision making

requirements of our language programs, the desires of our fellow teachers, and, above all, the needs of our students.

Decision making processes differ widely from program to program, and only the teachers and administrators involved in a program can determine what types of tests are needed. However, I have tried to outline some of the ways that testing and decision making are handled in the ELI at the University of Hawaii where testing has become more than just a "necessary evil." Tests have been incorporated into the ELI curriculum to perform a variety of necessary decision making functions. Since we want to make these decisions in the most responsible manner possible—precisely because we do care about our students—testing has become an integral and essential part of the language teaching that we provide.

**James Dean Brown** is director of the English Language Institute, Department of English as a second Language, University of Hawaii at Manoa. He is the author of *Understanding Research in Second Language Learning: A Teacher's Guide to Statistics and Research design* and numerous articles on the role of testing in language instruction.

### References

- ACTFL. (1986). *ACTFL proficiency guidelines*. Hastings-on-Hudson, NY: American Council on the Teaching of Foreign Languages.
- Alderson, J.C., Krahnke, K.J., & Stansfield, C.W. (1987). *Reviews of English language proficiency tests*. Washington, DC: TESOL.
- Bachman, L.F. (1989). The development and use of criterion-referenced tests of language proficiency in language program evaluation. In K. Johnson (Ed.), *The second language curriculum*. London: Cambridge University Press.
- Berk, R.A. (1980). *Criterion-referenced measurement: The state of the art*. Baltimore: Johns Hopkins University Press.
- Brown, J.D. (1981). Newly placed versus continuing students: comparing proficiency. In J.C. Fisher, M.A. Clarke & J. Schachter (Eds.), *On TESOL '80* (pp. 111-119). Washington, DC: TESOL.
- Brown, J.D. (1984a). Criterion-referenced language tests: What, how and why? *Gulf Area TESOL Bi-annual, I*, 32-34.
- Brown, J.D. (1984b). A cloze is a cloze is a cloze? In J. Handscombe, R.A. Orem, & B.P. Taylor (Eds.), *On TESOL '83* (pp. 109-119). Washington, DC: TESOL.
- Brown, J.D. with Christensen, T. (1987a). Interview: James D. Brown. *The Language Teacher, 11* (7), 6-10.

- Brown, J.D. (1987b). False beginners and false starters: How can we identify them? *The Language Teacher*, 11 (14), 9-11.
- Brown, J.D. (1988). *Understanding research in second language learning: A teacher's guide to statistics and research design*. London: Cambridge University Press.
- Brown, J.D. (1989a). Improving ESL placement tests using two perspectives. *TESOL Quarterly*, 23, 65-83.
- Brown, J.D. (1989b). Language program evaluation: A synthesis of existing possibilities. In K. Johnson (Ed.), *The second language curriculum* (pp. 222-241). Cambridge: Cambridge University Press.
- Brown, J.D. (1989c, March). *Short-cut estimates of criterion-referenced reliability*. Paper presented at the Language Testing Colloquium in San Antonio, Texas.
- Cziko, G.A. (1983). Psychometric and edumetric approaches to language testing. In J.W. Oller, Jr. (Ed.), *Issues in language testing research* (pp. 289-307). Rowley, MA: Newbury House.
- Davidson, F., Hudson, T., & Lynch, B. (1985). Language testing: Operationalization in classroom measurement and L2 research. In M. Celce-Murcia (Ed.), *Beyond basics: Issues and research in TESOL* (pp. 137-152). Rowley, MA: Newbury House.
- ETS. (1987). *Test of English as a foreign language*. Princeton, NJ: Educational Testing Service.
- Glaser, R. (1963). Instructional technology and the measurement of learning outcomes: Some questions. *American Psychologist*, 18, 519-521.
- Hudson, T. (1989). Mastery decisions in program evaluation. In K. Johnson (Ed.), *The second language curriculum* (pp. 259-269). Cambridge: Cambridge University Press.
- Hudson, T., & Lynch, B. (1984). A criterion-referenced approach to ESL achievement testing. *Language Testing*, 1, 171-201.
- Popham, W. J. (1978). *Criterion-referenced measurement*. Englewood Cliffs, NJ: Prentice-Hall.
- Popham, W. J. (1981). *Modern educational measurement*. Englewood Cliffs, NJ: Prentice-Hall.
- Popham, W.J., & Husek, T.R. (1969). Implications of criterion-referenced measurement. *Journal of Educational Measurement*, 6, 1-9.

## **BOOK REVIEWS**

**LANGUAGE TRANSFER: CROSS-LINGUISTIC INFLUENCE IN LANGUAGE LEARNING. Terence Odlin. Cambridge: Cambridge University Press, 1989. 210 pp. ¥2,360.**

Language transfer—or the influence of a previously acquired language, usually one's native language, on a second language—has long been a controversial topic among applied linguists and language teachers. In the heyday of contrastive analysis and audiolingualism in the 1950s, it was believed to be the major influence on second language acquisition (SLA). But in the '60s and '70s, with the Chomskyan revolution in linguistics and the emphasis on universal acquisition processes, and as it became clear that contrastive analysis by itself could not predict learner errors, many researchers lost interest in language transfer.

However, the 1980s saw a resurgence of interest in the phenomenon of transfer, as is evidenced by numerous journal articles, and by major collections of research papers by, for example, Gass and Selinker (1983) and Kellerman and Sharwood Smith (1986). Terence Odlin's *Language Transfer* is an ambitious attempt to integrate and interpret the research in a way that is in step with current SLA thinking. The author should be commended for producing a readable introduction to such a complex field.

Odlin gives a balanced overview of the study of language transfer from the perspectives of its historical roots, of problems in the study of transfer, of linguistic subsystems, and of nonstructural factors related to transfer. In contrast to the traditional view represented by Lado (1957), which portrayed native language influence as the central issue in language learning, transfer is now considered as but one factor—albeit an important one—in SLA. Language transfer, of both the positive and negative varieties, is seen to interact with universal processes and language typology, as well as with a complicated array of nonstructural factors, such as personality, age, and social context, in shaping second-language performance.

*Language Transfer* consists of ten chapters, each of which is divided into numbered sections. Throughout the book, consider-

able cross-referencing of sections has helped to integrate the material and to make it possible for readers to concentrate on topics that are of particular interest to them. In addition, each chapter begins with a clear introduction outlining the material to be covered. The author avoids the use of overly technical terms where possible, and those that are used have been put in a glossary at the end, where reference is made to the sections in which the terms first occur. There is also a very extensive bibliography. These features make the book suitable as an introduction to the field for newcomers, as well as a reference work for experienced researchers in language transfer or other areas of SLA. Nevertheless, certain parts may still prove difficult for the reader who has no background in linguistics.

The first three chapters introduce the phenomenon of language transfer and locate it within the context of SLA and linguistic research as a whole. In Chapter 1, the author notes the primary aim of the book is "to reconsider the problem of transfer in light of recent second language research" (p. 4). He outlines reasons why language teachers and linguists should consider more closely the problem of transfer. Teachers should consider it, Odlin maintains, in order to pinpoint the problems of particular groups of students, and to understand why certain aspects of the target language are difficult or easy for learners. In Chapter 2, a historical overview of the debate over the role of transfer is given. The controversy is traced to the nineteenth-century historical linguists whose main interest was not SLA or language teaching, but language classification and language change; that is, to the debate over whether language change can best be explained by a "tree model," where primacy is accorded to internal development, or by a "wave model," where linguistic borrowing and language mixing are seen as major influences. This debate, of course, parallels that between the contrastive analysts of the 1950s (whose major concern was cross-linguistic influence) and the universalists of the '60s and '70s (whose concern was internal development). Chapter 3 discusses four types of theoretical and practical problems in the study of language transfer: problems of definition (of what constitutes transfer), of comparison (across language systems), of prediction (of when learner errors will occur), and of generalization (in

relation to language universals and linguistic typologies). From these first three chapters emerges the author's perspective that neither an extreme universalist view of language learning, nor a view based exclusively on language transfer, can be supported, and that there is a need for a more balanced type of thinking.

The next four chapters examine empirical studies of transfer in relation to specific linguistic subsystems: discourse, semantics, syntax, and phonology and writing systems. The author makes it clear that this classification is somewhat arbitrary, and that there is considerable overlap among the subsystems. In all areas, however, the evidence points to the importance of language transfer in language learning. But, as this often operates together with factors of language typology and universal processes, it is sometimes necessary to look well below the surface to discover the exact role of transfer.

Chapter 4 may be of special interest to many *JALT Journal* readers, as it discusses areas in which one finds great differences between Japanese and Western languages, and where there is great potential for negative transfer and miscommunication in language learning. This chapter, "Discourse," actually deals with a wide range of studies, including those that many researchers would label as "sociolinguistics" or "pragmatics." The author's survey of research here focuses on two general areas: politeness, including requests, apologies, and conversational style; and coherence, including narrative structure and indirection. In both of these areas, many of the studies presented deal with differences between the Japanese and English discourse systems, and with related communication problems of Japanese learners of English.

One shortcoming of this section is that many of the studies cited point merely to differences between discourse systems (as did early contrastive analyses with phonology and syntax), and do not look at actual learner behavior. However, this can be seen as a reflection of the state of the art in relation to discourse transfer, rather than of the author's approach.

Chapters five to eight review research in semantics, syntax, phonology, and writing systems, respectively. Chapter 5 looks at the role of transfer in propositional semantics (which is closely related to syntax in "if-then" statements, for example, and active

vs. passive voices), as well as in the better investigated area of lexical semantics (e.g., transfer of vocabulary, metaphor, and morphological endings). Chapter 6 concentrates on three areas of syntax where the number of existing studies of both positive and negative transfer is large: word order, relative clauses, and negation. It is in such syntactic studies that the debate over the role of language transfer has been (and continues to be) most vocal, but it is also here where many of the attempts to reconcile language transfer with universal processes have been made. Chapter 7 looks at phonetics and phonology, the area where the role of native language influence is probably the least controversial. In this chapter the author also briefly examines how transfer affects one's performance in a foreign writing system, noting that there is a close connection between language orthography and native language pronunciation.

Chapter 8 is unique in its consideration of factors that are usually not discussed in contrastive studies: that is, "nonstructural" factors. These include personality, proficiency level, literacy, age of acquisition, and social context, among others. Many of these factors are related to individual variation in language learning, and hence pose problems for the researcher trying to describe the collective behavior of members of a linguistic community. In discussing these, the author draws heavily on the research on child bilingualism and on language use in multilingual communities, as well as citing research more directly related to SLA. This chapter should thus be of interest to a wider range of readers than some of the others.

Chapter 9 pulls together the research discussed in the previous chapters and draws out some common themes. Among these are the conclusions that transfer occurs in all linguistic subsystems, that it occurs in both informal and formal contexts, that it occurs among children as well as adults, and that similarities between languages can facilitate learning. A number of areas are then suggested for further research. These include social context, longitudinal comparisons, acquisition of non-European languages, and child bilingualism.

Chapter 10, which considers implications of the research for teaching, might on the surface appear to be the chapter with the

greatest relevance to *JALT Journal* readers. The author's conclusions are decidedly cautious, however, and rightly so, considering the complexity of the phenomenon of transfer and the incomplete nature of the research. Teachers who go into the chapter with the expectation of getting concrete suggestions on how to conduct their classes will be disappointed. The major implication here is that teachers should be aware of the existence of native-language influences in second language learning and that they should take an interest in the language and culture of their students. For most of us, that is certainly nothing new.

Seen as a whole, however, *Language Transfer* holds another, more implicit, message for language teachers. While the research may provide few clues as to how one *should* teach a class, it provides strong clues as to how one *should not* teach. In the same way that research on language universals discredited certain teaching methods, notably audiolingualism, that artificially maximized the role of native-language influence in language learning, so should the research Odlin describes discredit methods that minimize the role of transfer in favor of strong interpretations of linguistic innateness (e.g., the Natural Approach [Krashen and Terrell, 1983]). There is a lot of middle ground between the extremes of native-language influence and linguistic innateness, and it is the language teacher's responsibility to explore—and exploit—this territory.

**Reviewed by Fred E. Anderson, Fukuoka University of Education**

### References

- Gass, S., & Selinker, L. (Eds.). (1983). *Language transfer in language learning*. Rowley, MA: Newbury House.
- Kellerman, E., & M. Sharwood Smith (Eds.). (1986). *Crosslinguistic influence in second language acquisition*. New York: Pergamon.
- Krashen, S.D., & Terrell, T.D. (1983). *The natural approach: Language acquisition in the classroom*. Hayward, CA: The Alemany Press.
- Lado, R. (1957). *Linguistics across cultures*. Ann Arbor: University of Michigan Press.

**BILINGUALITY AND BILINGUALISM. Josiane F. Hamers and Michel H.A. Blanc. Cambridge University Press, 1989. 324 pp. ¥11,900 and ¥4,400.**

Since 1953, when U. Weinrich wrote his classic study, *Languages in Contact*, scholars have moved from

the conception of the bilingual as someone whose speech shows interferences and who is at worst cognitively deficient and at best the sum of two monolinguals...to the conception of an integrated person...for whom bilingual experience may enhance cognitive functioning (p. 257).

*Bilinguality and Bilingualism* not only traces the history of this transition, but also attempts to bring the reader up to date on what the authors call the "state of the art" about bilingualism. In what could be viewed as a "textbook" on the subject of languages in contact, Hamers and Blanc try to introduce all major current theories on the subject, evaluate the experimental evidence to sustain or refute these theories, and indicate areas where further inquiry is required.

The range of scholarship covered in this single volume is extensive. In dealing with bilingualism on the individual, interpersonal, and intergroup levels, the authors adopt a multidisciplinary approach that includes linguistics, psychology, neurology, sociology, and education. Bilingual (French and English) and bicultural themselves, the authors are well equipped to handle the task.

The first six of the book's ten chapters deal with bilingualism on the personal level—what the authors call "bilinguality." The authors begin their study of this phenomenon by discussing the various definitions that have been made of bilinguality, ranging "from a native-like competence in two languages to a minimal proficiency in a second language" (p. 7). They then move on to discuss the systems used to measure bilinguality, concluding that all are inadequate.

Chapter 2, "The Ontogenesis of Bilinguality," notes that early research on bilinguality was conflicting: documented child biographies recorded the "harmonious development of the bilingual child," while early psychometric studies indicated developmental

delay in bilingual children, compared to monolingual peers (p. 31). In trying to determine the cause of this divergence, the authors discuss current research on the development of bilingual children, linguistic mixing, the neuropsychological development of bilinguals, and lateralization of the brain in monolinguals. They also discuss various types of bilinguals, bilinguality, cognitive development, and the role of sociocultural environment on cognitive development in bilinguals. Other topics include infant bilinguality, consecutive bilinguality, endogenous and exogenous bilinguality, additive and subtractive bilinguality, "semilingualism," and code switching and mixing.

By the end of this chapter, the authors have articulated one of the major themes of their work: the difference between additive and subtractive bilinguality. They show that, whereas bilinguality can actually enhance cognitive functioning when both languages are valorized, it can hinder development when a child's first language is not valorized by society at large. Frequently, a second language is used to introduce literacy skills before the child has been prepared for them in his or her first language. This means that subtractive bilinguality is most often a problem of children from minority groups. The authors therefore stress measures that can be taken by educators to prevent this phenomenon in minority children.

This theme is expanded in the book's third chapter, "The Social and Psychological Foundations of Bilinguality." The first two sections of this chapter deal with the development of language in a social context in general, while the third section relates this knowledge to bilinguality.

Chapter 4 tries to develop a model for the workings of the bilingual mind by comparing it to a computer. The terminology here is that of the age of electronics: the authors refer to switching, on-line processing, input, output, memory, storage, and retrieval, though they do not completely accept any current model for the functioning of the bilingual brain.

Chapter 5 begins a transition from the micrological to the macrological level, as the discussion turns to the sociopsychological concepts of culture and identity in the bilingual. The authors move from the individual to interpersonal relations and finally to inter-

group relations, and discuss the effect of group identity on speech patterns, including speech accommodation. They then move on to the way group affiliation effects personal perception of language. The view of languages in contact here takes on darker shades, as we see interethnic strife reflected in linguistic boundaries.

The authors point out that "As with the notion of ethnic group identity (which they had thoroughly discussed), language identity is very much a function of the interlocutor's perceptions" (p. 157). In fact, "what is defined by linguists as one and the same language...may in fact be perceived as different languages by different speakers of that language." They go on to give the example of Hindustani, which is viewed as three different languages because of the different cultural, religious, and political allegiances of the main groups that speak it. And, in what may seem a particularly relevant insight in these days of "language revolutions" in the Soviet Union, the authors note, "Contemporary history abounds in examples of nationalist movements based essentially, though not exclusively, on language demands" (p. 160).

This lengthy discussion of the political ramifications of languages in contact is followed by a chapter on bilingual education solidly based on the lessons of intergroup relations that preceded it. In addition to discussing the various problems of national language planning, this chapter clearly differentiates between the requirements of bilingual education for dominant groups and those for ethnic minority children.

This wide-ranging work concludes with chapters on second language acquisition and translating and interpreting. The former does not discuss teaching methodology, but instead describes the psychological and linguistic principles of bilingual development upon which second language education should be based. It stresses that "there is no simple and easy application of theories of L2 acquisition to L2 teaching" (p. 243), and indeed this chapter seems too abstract for the average foreign language teacher. The final chapter, too, while dealing with a fascinating subject, seems tangential to the major thrust of the work.

As can be seen from this brief outline of its contents, *Bilinguality and Bilingualism* integrates a vast amount of research. Its 42-page reference list—in itself a valuable resource for anyone interested in

bilingualism—includes 851 publications in English and French, the vast majority in English. It is current, with the latest entries published in 1988. Moreover, the cited research comes from fields as diverse as neurology and history, and covers an impressive array of countries and languages.

Although at times descriptions of research are obscured by highly technical terms, the authors are able to deliver balanced analyses of findings. They not only offer an excellent survey of what theories have been proven, but they also provide insightful suggestions as to what remains to be done.

For the reader without a solid background in the field, however, it would be helpful to have explanations of such research methods as longitudinal observation, evoked potential measures, and cross-language flanker technique. The reader unfamiliar with French, German, and Spanish may have trouble understanding descriptions of some experiments involving those languages. Section titles could also be more helpful. For example, the section discussing models for L2 acquisition is entitled "Contrastive analysis, error analysis and interlanguages."

Finally, the section on code alternation, while describing some very interesting studies, remains on too abstract a level to give much insight into this phenomenon. The authors are content to state that, like "code-switching, code mixing can be a bilingual's specific code which enables him to express attitudes, intentions, roles, and to identify with a particular group" (p. 152). Concrete examples of how bilinguals use code alternation to quote, to discuss certain topics, and most importantly, to personalize discourse or create dramatic effects that intensify narratives, would have given more life to the authors' argument that "code-switching can become an autonomous code..." (p.154).

Such faults should not detract from the positive contributions of this work: the volume provides a thorough description of what is known about a very complex subject that has only relatively recently received academic attention.

The authors state that the book is "meant for all those who are interested in language behavior" (p. 4). For many, however, reading selected sections might prove more useful than a cover-to-cover study. For example, parents striving to raise children bilingually

might find great comfort in Chapter 2, Section 3, dealing with bilinguality and cognitive development. Similarly, teachers and those who form education policy in multicultural areas could obtain a clear set of principles on how to prevent subtractive bilinguality in educating minority children (Chapter 8). For any reader, though, difficulties in reading *Bilinguality and Bilingualism* should prove well worth the struggle.

**Reviewed by Mary Goebel Noguchi, Doshisha University  
and Doshisha Women's Junior College.**

**ESP IN THE CLASSROOM: PRACTICE AND EVALUATION.** Dick Chamberlain and Robert J. Baumgardner (Eds.). *ELT Documents: 128, Modern English Publications, 1988. 165 pp.*

In a world where advances in science and technology are being made daily, business is being conducted on multi-national levels, and mass communications are bringing people of all nations together, English has emerged as the medium of exchange. The need for professionals trained in English has become a major concern. The teaching of English for Specific Purposes (ESP) has arisen in an attempt to train professionals in the kind of English used in their domain of expertise.

Teachers in the field of ESL/EFL who are required to teach ESP are always on the lookout for resource materials which will give them practical and useful approaches to developing syllabi and course materials. *ESP in the Classroom: Practice and Evaluation* promises to be such a resource guide.

*ESP in the Classroom* is an anthology of articles originally presented as papers at the International Conference on English for Specific Purposes, held in Sri Lanka in 1985. The book is divided into five sections of three articles each. In their attempt to achieve this symmetry, the editors have at times placed an article in one section when it really belongs in another.

The first section, "Viewpoints," begins with the keynote address by the famous scientist and author Arthur C. Clarke, who lives in Sri Lanka. He notes the importance of English for electronic communications—gaining access to computer data bases, exchanging electronic mail, and so forth. The second article, by Kachru, although rather long and convoluted, points out the specialized needs of ESP at the intranational level. He makes a case for regional varieties of English by drawing a distinction between communications meant for the "inner" versus the "outer" circle. The last article in this section, by Dharmapriya, describes historically the rise of English during the late 1940s in Sri Lanka, and its subsequent decline. He emphasizes, however, that the use of English for science and technology, due to the lack of native language resources and reference materials, has again enhanced

the importance of English in Sri Lanka, as well as in other third world nations.

The second section deals with the teacher and the learner. Strevens makes a distinction within the field of ESP between English for Academic Purposes (EAP) and English for Occupational Purposes (EOP). He states that learners in EAP are usually younger and less goal-oriented, whereas the EOP learners are often older, having worked in a field of specialization, and therefore have specific language objectives. Strevens sees ESP as a professional challenge which may result in general English teachers placing themselves in the new role of "enabler."

Ramani also advocates a "user-based" approach in which task-oriented, problem-solving materials are used. She describes her work with doctoral students in engineering who need English to publish their research findings. She stresses the importance of analyzing the discourse of journal articles; then, with the use of peer interaction, the students produce their own research findings in publishable form. The last article in this section, by Chamberlain, is a rather esoteric discussion about the role of the teacher. Like Strevens's paper, it highlights the teacher's role as language moderator.

The third section attempts to bridge the gap between the classroom and the world of work with a discussion of issues in methodology. Huckin writes that ESP's role is to develop a professional level of communicative competence in English within the frame of a four-year technical education. English courses in the first two years should follow a traditional method of technical textbook analysis and lecture comprehension exercises. In the final two years, the focus should shift to a "case study" method, which would require the development of problem-solving skills and technically appropriate argumentation. Team-teaching with faculty from the technical fields forms a vital component of Huckin's case study approach. Hutchison advocates a process approach to learning whereby students are challenged through the use of task-based assignments to increase their linguistic level to match their conceptual level. Mountford's article deals with materials rather than methods, so it seems misplaced in this section of the anthology. He criticizes available ESP materials as being either too subject

specific or too skill specific. He supports the production of teacher-prepared materials which would fit the needs and interests of the students.

The fourth section of the book addresses issues in the evaluation and administration of ESP programs. Alderson's rather lengthy article gives a good overview of the principles of testing in general, without every targeting the specifics of testing ESP. Pearson's article also is not ESP specific. He writes about the "washback effect" of testing. Teachers teach for exams. If the exams are proficiency based, the teachers will motivate the students toward this end. Tan's article describes in detail the English service department of a large English-medium polytechnic institute in Singapore. Especially interesting is her description of the coordination between the English department and the technical departments.

The final section of *ESP in the Classroom* raises issues in designing and implementing materials. The first article, by Cumaranatunge, does not in fact deal with materials, but describes the process of assessing the English language needs of domestic workers in West Asia. Baumgardner and Tongue advocate the dubious use of contrastive analysis to demonstrate standard English forms. They suggest contrasting non-standard English newspapers (such as can be found in Sri Lanka) with a native-speaker model. This article seems to be in direct opposition to the views held by Kachru (section one). The final article in this anthology, by Baumgardner et al., presents a case study describing materials used in a two-year engineering program in Sri Lanka.

As with any anthology, *ESP in the Classroom* offers some useful articles, while others lack merit for an ESP source book. As ESP is a relatively new field, there is much theory without much practical guidance for the classroom teacher. The title *ESP in the Classroom: Practice and Evaluation* promises more than most of the articles in this volume have to offer.

**Reviewed by Sarah Rilling, Science University of Tokyo.**

## Information for Contributors

### EDITORIAL POLICY

*JALT Journal* welcomes practical and theoretical articles concerned with foreign language teaching and learning in Japanese, Asian, and international contexts. 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 articles and reports, reviews, and comments on earlier *JALT Journal* writings (for the "Point to Point" section). Articles should be written with a general audience of language educators in mind. Statistical techniques and unfamiliar terms should be 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, D.C.). Consult recent copies of *JALT Journal* or *TESOL Quarterly* for examples of documentation and reference lists. **This is a strict requirement.** Also, remember to give precise page numbers of cited work in both the text and reference list.

#### Format

No longer than 20 pages, including reference list, typed on A4 or 8 1/2" X 11" paper, and double-spaced. Writers must supply camera-ready diagrams or figures (if any) before final publication.

#### Materials to be Submitted

- Two paper copies of the manuscript, plus one Mac, IBM PC, or NEC 9801 computer disk version (to be returned; please use tab settings to indent and to make columns), or
- Three copies of the manuscript

- Abstract (less than 200 words)
- Japanese translation of title and abstract if possible
- Running head title (1 to 4 words)
- Biographical sketch(es) (no more than 50 words)

### **Evaluation Procedure**

Manuscripts are subject to blind review by two readers. The author's name and references that identify the author should appear only on the cover sheet. Evaluation is usually completed within two months.

### **Restrictions**

Papers sent to *JALT Journal* should not have been previously published, nor should they be under consideration for publication elsewhere. We regret that paper manuscripts cannot be returned.

### **Address for Manuscripts and Inquiries**

Mr. Malcolm J. Benson  
*JALT Journal* Co-Editor  
 Hiroshima Shudo University  
 1717 Ohtsuka, Numata-cho  
 Asaminami-ku, Hiroshima 731-31

### **Reviews and Address for Reviews**

We invite submissions of reviews of all 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*. Contact the reviews editor for guidelines and send all review materials to her at the following address:

Ms. Jane Wieman  
*JALT Journal* Reviews Editor  
 11-7 Miyanomae-cho Arashiyama  
 Nishikyo-ku, Kyoto 616

### **Address for Inquiries about Advertising**

Advertising  
*JALT* Central Office  
 Lions Mansion Kawaramachi #111  
 Kawaramachi Matsubara-Agaru  
 Shimogyo-ku, Kyoto 600

---

JALT JOURNAL 第12巻 第1号

平成2年 4月25日 印刷

平成2年 5月 1日 発行

発行人 全国語学教育学会 (JALT) 事務局長  
THOMAS N. ROBB

発行所 JALT 本部  
〒600 京都市下京区寺町松原上ル  
ライオンズマンション111号  
TEL (075) 361 - 5428

印刷所 厚 進 社  
〒530 大阪市北区天満2丁目13-3  
TEL (06) 351 - 8795

---