

昭和五十四年四月二十一日 第三種郵便物認可

ISSN 0287-2420

¥750

# JALT JOURNAL

JOURNAL OF THE JAPAN ASSOCIATION  
OF LANGUAGE TEACHERS

全国語学教育学会

VOLUME 14 NO. 1

MAY 1992

**CONTENTS**

<b>Editorial</b> .....	4
------------------------	---

**Articles**

Testing Listening Comprehension Using Summary Cloze <i>Jo A. Lewkowicz</i> .....	7
Using Learner Evaluations to Improve Language Teaching <i>Keith Maurice</i> .....	25
EFL Learners Talking to Each Other: Task Types and the Quality of Output <i>Atsuko Ushimaru</i> .....	53

**Point-to-Point**

A Response to "Communicative Competence and the Japanese Learner" <i>Ronald Sheen</i> .....	69
Rod Ellis Replies .....	74

**Book Reviews**

An Introduction to the Pronunciation of English (A. C. Gimson) English Phonetics and Phonology: A Practical Course (Peter Roach) <i>Reviewed by Tim Riney</i> .....	77
Conditions for Second Language Learning (Bernard Spolsky) <i>Reviewed by Richard J. Marshall</i> .....	83
Second Language Writing: Research Insights for the Classroom (Barbara Kroll) <i>Reviewed by Sandra S. Ishikawa</i> .....	89
Currents of Change in English Language Teaching (Richard Rossner and Ron Bolitho) <i>Reviewed by Ronald Sheen</i> .....	93

<b>Information for Contributors</b> .....	98
---	----

**Editors**

**Malcolm Benson**  
*Hiroshima Shudo University*

**Charles Wordell**  
*Nanzan University, Nagoya*

**Reviews Editor**

**Roger Davies**  
*Nagasaki Prefectural University*

**Japanese Language Editor**

**Kyoko Nozaki**  
*Kyoto Sangyo University*

**Editorial Board**

**Charles E. Adamson, Jr.**  
*Shizuoka Institute of Science and Technology*

**John Fanselow**  
*Columbia University*

**Michael Horne**  
*Nagoya University*

**Ann Johns**  
*San Diego State University*

**Susan Johnston**  
*Temple University Japan*

**Teruyuki Kume**  
*Kanda University of International Studies*

**Michael Liggett**  
*Hosei University, Tokyo*

**John Maher**  
*International Christian University*

**Peter McCagg**  
*International Christian University*

**Joy Reid**  
*University of Wyoming*

**Lawrence Schourup**  
*Osaka Women's University*

**John Shillaw**  
*Kyoto Institute of Technology*

**Publications Chairperson: Carol Rinnert**

***The Language Teacher* Editors:**

**Carol Rinnert, Greta Gorsuch**

**Layout: Charles Wordell**

# **JALT NATIONAL OFFICERS 1992**

## **President**

**Philip Crompton**

**Treasurer**  
**Aleda Krause**

**Recording Secretary**  
**Steve Mason**

**Programs**  
**Carl Adams**

**Membership**  
**Setsuko Toyama**

**Public Relations**  
**Masaki Oda**

## **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**

## Editorial

**Jo A. Lewkowicz**, in her paper “Testing Listening Comprehension Using a Listening Summary Cloze,” demonstrates how a contextualized test of listening comprehension can provide teachers and students with a positive washback effect. As a result of pilot work in China and Hong Kong, the “listening summary cloze” test is seen as allowing teachers great flexibility regarding the material to be used in the test, offering a good spread of marks, and most importantly encouraging students to develop meaningful habits in both listening and note-taking.

**Keith Maurice** looks at some ways in which the experiences and reactions of learners may be called upon and exploited to improve language-teaching programs. Case studies, drawn largely from the author's experience in Japan, show how learner feedback and administrative action can combine to provide the teacher with ideas for improving classroom effectiveness—though these ideas are not always taken in the spirit in which they are offered! The article comes at a time when teacher performance in the classroom is coming under closer and closer scrutiny, and when the traditional teacher-student relationship is being re-defined worldwide.

**Atsuko Ushimaru's** research paper examines learner output in a Japanese setting, showing how grammaticality is affected by the nature of the task. Subjects were given three tasks to complete, and progress towards a target native-speaker norm was recorded. The results are of interest in that they demonstrate how the most precise task produced the greatest number of errors, and vice versa. The effect of pressure on creating an utterance in a foreign language is therefore seen as significant, a finding which accords both with earlier research and with common sense.

In a Point-to-Point exchange, **Ronald Sheen** comments on Rod Ellis's article in the last issue of the *JALT Journal*, and **Rod Ellis** replies.

The issue concludes with four reviews. The first, by **Tim Riney**, compares two major texts dealing with English sounds, by A. C. Gimson and Peter Roach. **Richard J. Marshall** reviews Bernard Spol-

sky's theoretical model of the conditions for language learning, and Sandra S. Ishikawa describes a collection of essays dealing with second language writing. The final review, by Ronald Sheen, summarizes major articles drawn from issues of the *ELT Journal* published between 1982 and 1988 and now published in book form.

### Correction

In Kazuko Matsumoto's article in the November 1991 issue (Vol. 13, No. 2), there were several numbers misprinted in the table on page 150. The items in boldface italics are the corrections:

not	0	<i>0</i>	3
Copula Negation	'm not	isn't	wasn't
'm not	2	<i>0</i>	0
isn't	0	4	2
wasn't	0	<i>0</i>	1

### Appreciations

The editors thank two members who are leaving us—Larry Schourup, who has served as a member of the Editorial Board since May 1988, and Jane Wieman, who has been our Reviews Editor since November 1988. We also wish to thank Michael Rost for help on the current issue.

### In Memory: Roger Griffiths

We deeply regret the accidental death, while swimming in Thailand on August 2, 1991, of Roger Griffiths. Roger contributed two excellent articles to the *JALT Journal* and assisted us in the evaluation of submissions related to L2 learner comprehension of spoken input. An imaginative and prolific scholar, Roger will long be missed by his colleagues throughout the world.

**CAMBRIDGE** ELT: the professional choice

## Evaluating Second Language Education

**NEW**

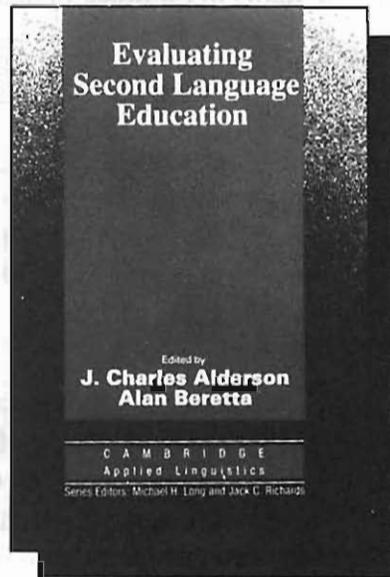
Edited by J. Charles Alderson and Alan Beretta

*Evaluating Second Language Education* examines important theoretical and methodological issues in the evaluation of second language education.

It offers:

- a range of state-of-the-art case-studies
- a review of the literature showing past developments
- useful input to teacher training programmes.

This book will be a valuable resource both for those already engaged in evaluation and for those in training.



For further information, please contact:  
David Fisher, Cambridge University Press,  
United Publishers Services Ltd, Kenkyu-sha  
Building, 9 Kanda Surugadai 2-chome,  
Chiyoda-ku, Tokyo 101.  
TELEPHONE: (03) 3295-5875

 **CAMBRIDGE**  
UNIVERSITY PRESS

## Testing Listening Comprehension Using Listening Summary Cloze

Jo A. Lewkowitz

*Language Centre, The University of Hong Kong*

This paper discusses the development of a listening summary cloze test which was initially devised in response to demands from tertiary-level teachers in China for a notetaking and listening test that would have a positive washback effect on their teaching. The paper first considers the theoretical implications underlying the development of an alternative testing format and then describes the process of piloting the new test. The pilot was carried out in two phases: initially in an EFL situation in China, and later in an ESL situation in Hong Kong.

### Listening Summary Cloze を使用しての聴解力テスト

当稿は最初中国の大学教師よりの要求に応じて彼らの教授法に対して明確な washback effect が得られるであろうノートの要約と聞き取りテストの為の listening summary cloze test の開発を検討するものである。最初に他のテスト様式 (format) の基本的発展に基づいた理論的考察を行い、次に試験的な新テストのプロセスを述べる。

この試験は最初に中国の EFL のプログラムで、次にホンコンの ESL のプログラムで実施された。

### 1. Introduction

The testing of listening comprehension has a comparatively short history. Prior to developments in technology and the audio-lingual method of teaching, listening tests were limited to tests of dictation. In the last fifteen to twenty years, however, listening tests have not only become much more widespread, but also more varied. Nowadays, it is not uncommon for listening comprehension tests to be made up of a battery of subtests, testing a cross-section of listening skills.

Over the years, changes in test tasks used in large-scale tests have tried to reflect our ever-increasing understanding of the processes involved in listening. When tests of auditory discrimination were introduced, it was commonly believed that the first stage in the understanding of a message was that of understanding the sounds. In other words, it was thought that listening was a one-way, bottom-up process starting with the lower-order skills such as recognising phonemes and morphemes, and ending with the higher-order skills such as inference. Furthermore, it was believed that "if a learner was tested in phoneme discrimination, stress and intonation, the sum of the 'discrete' sub-tests would be equivalent to his proficiency in listening comprehension" (Weir, 1988, p. 55).

It is now thought that listening is a much more complex process which involves simultaneous top-down and bottom-up processing to which no fixed order can be ascribed. To comprehend a message, listeners not only use this knowledge of the sound system, lexis, and syntax of the language, but also their knowledge of the outside world and how it relates to the topic at hand, as well as their knowledge of the culture of the language. In addition, they have to utilize their understanding of what has been said so far to anticipate what will follow. Thus, with a growing acceptance that context plays a very important part in the understanding of spoken discourse, it is not surprising that contextualized discourse is increasingly being used as the basis of listening comprehension tests. Learners are being required to listen to complete utterances, dialogues, and monologues in the form of, for example, mini-lectures, instead of to uncontextualized words or short exchanges. Tests appear to this extent to have caught up with the theory of listening and to reflect demands similar to those placed on learners in the real world.

One result of this change is that an increasing number of listening tests have become more integrative in nature, testing more than one subskill or language element at a time. However, this is not true of all listening tests. Where teachers are responsible for setting their own tests, they often prefer to use test formats that have been tried and

## Listening Summary Cloze Tests

proven to give statistically reliable results than to try out new test types. Hence, they adhere to the discrete point test based on uncontextualized discourse rather than develop a new test in line with current theory.

### 2. Difficulties of Setting Contextualized Tests

Developing contextualized listening tests is a complex and time-consuming task. Fundamental is the construction and recording of suitable texts. Relevant topics need to be identified and either suitable prerecordings selected or custom-made recordings arranged. Prerecordings, despite having the advantage of being authentic, are often difficult to adapt for testing purposes; they may, for example, be insufficiently clear, too long, or unnecessarily complex. Custom-made recordings, on the other hand, may sound stilted. It is, for example, very difficult to record a written talk or lecture in a manner and at a speed that sounds authentic.

Another area of difficulty for the test writer is that of identifying what is to be tested. When it was believed that the sum of the discrete parts made up the whole, it was possible to sample a range of parts and claim that by testing sound discrimination, recognition of stress, and of intonation, a picture of an individual learner's listening skills could be drawn up. However, with the onset of integrative tests two problems emerged. First, it became impossible to separate the discrete parts since more than one subskill is tested at a time. Second, and perhaps more important, it became more difficult to isolate the skill of listening from the other language skills, since not only are the subskills of listening being integrated but so are the other language skills. The listening test that requires learners to respond in speech or writing is impure in that it tests more than one language skill at a time and thus raises questions of test validity and reliability, even though the tasks required of the learners may be more authentic in nature. As Rost (1990) points out:

The question of reliability and the issue of construct validity present an uncomfortable dilemma for test writers and language teachers. In order to include testing formats that allow for improved face validity and broader content validity, reliability may

be compromised. Specifically, if listening is to be measured in situations of actual use, situations in which intervening variables obviously come into play, reliability of measurement will be compromised. However, it is precisely the integration of information-gathering skills from spoken texts with other pragmatic skills that serves as the basis for listening development. While it is important for pedagogic reasons to identify learner problems with listening, at the same time it compromises educational principles to equate listening ability with information-gathering skill. (p. 180)

Related to the above are problems of test format. The reason discrete point tests have been found to be statistically reliable is they have usually taken the form of objective, true/false or multiple-choice items based on single words or short exchanges, with the tester being able to ask a large number of questions. This type of test has led teachers to base their teaching on low-level recognition exercises such as appear on traditional tests. Such tests, by restricting the practice learners get in understanding longer stretches of discourse, have had a negative washback effect on teaching. If students are to understand spoken language outside the classroom setting, they need to be exposed to samples of such language. They therefore need to be tested on their understanding of discourse which varies in length, mode, and purpose. For the test to have a positive washback effect on teaching, learners need to be asked to do more than merely answer true/false or multiple-choice type questions. They need to produce language in response to what they have heard. In other words, as Bachman (1990) points out, "language is both the object and the instrument of our measurement," and hence "it is extremely difficult to distinguish the language abilities we want to measure from the method-facets used to elicit language" (pp. 287-288).

Contextualization, therefore, makes more difficult not only the selection of the text, but also the setting of the questions that assess students' comprehension of it. This may be the reason why many teachers shy away from developing such tests. They frequently ques-

## Listening Summary Cloze Tests

tion their own competence in the field of testing, and thus Buck (1988), in the context of Japanese university entrance examinations, comments that “teachers see little alternative but to produce a different version of the tests made in previous years, or imitate the tests of more prestigious institutions” (p. 17). This phenomenon is, of course, also a result of other constraints, such as the time pressures put on teachers to write tests and the difficulty of validating new measures, as well as the scale of the testing procedure with which teachers have to cope.

### 3. The Use of Cloze for Testing Listening

The cloze test, first introduced by Taylor (1953) as a means of assessing the reading ability of native speakers, has since been adapted for use with non-native speakers. It has moved away from the principle of *n*th word deletions, that is true cloze as advocated by Taylor, to selective rational deletion procedures, sometimes referred to as modified cloze or gap-filling. This latter procedure allows testers to select appropriate deletions so as to test not only lower-order skills, but also a range of higher-order skills.

Further changes have been made in the techniques used for the questions and the marking procedure. Originally, every *n*th word was left blank and the testees were required to provide the exact word missing from the text. Adaptations have seen the development of the multiple-choice cloze as well as a more flexible marking scheme, whereby acceptable words rather than only the exact word are marked as correct.

Cloze has also been extended to the testing of listening. In such tests the students are required to fill in the blanks in a written passage based on information they have heard. To avoid the completion of these blanks being done without listening, testers have found it necessary to ensure that the deletions are made on high-information content words or phrases. Yet even this does not guarantee that students are listening to more than just individual words missing from the text. Therefore, modifications have been tried whereby students are given a cloze summary of the original either in the L1 or L2. An example of such a

test is given in Heaton (1983, pp. 82-83) and Buck (1988) describes how this method has been developed successfully in Japan.

#### **4. The Development of the Listening Summary Cloze**

The present study is concerned with an extension of the listening cloze test described above. The test's main difference is that it is made up of two tasks rather than one, and though this format has been described elsewhere (Heaton, 1983, pp. 83-85), it appears not to have been used with summary completion tasks. The first task for the students in this study was that of notetaking, which the students did while listening to a mini-lecture, while the second task, the only one that was assessed, required the students to complete a summary cloze passage with the aid of their notes. The purpose of the study was to see how well a test of this nature performed in comparison with the more traditional type of listening comprehension tests both in terms of reliability as well as validity.

This test, the listening summary cloze, was developed for use with tertiary level students in a foreign language environment in China. It was designed with two objectives in mind: (a) The test was to measure students' ability to listen to longer stretches of discourse such as lectures, and (b) it was to encourage good notetaking both during the test as well as in class. In other words, the test was specifically designed to have a positive washback effect on teaching.

It was felt that notetaking was an important skill for the students to develop and, therefore, had to be tested in some way. Yet a review of the alternative methods of testing notetaking revealed that they were either too cumbersome to administer or would not have the desired washback effect on teaching. For example, having the students write a summary of a lecture was seen as very difficult to assess and the students' success in the task would have been dependent to a large extent on their writing ability. To make the task authentic, the students had to be required to use their notes for a particular purpose, as they would have to in their studies. Therefore the task developed, that of completing a summary of a text through gap-filling, was one that

## Listening Summary Cloze Tests

limited the amount of writing required and at the same time allowed for objective assessment, since only the summary completion, not the notes taken while listening to the lecture, was assessed.

### 5. Methodology

#### 5.1 Subjects

In the first instance, two versions of the listening summary cloze were piloted at the Guangzhou Institute of Foreign Languages in 1989, as part of a larger battery of proficiency tests.<sup>1</sup> The first, designed for learners majoring in English who had completed their second year at university (GEM 4), was administered to a total of 280 students—173 end of first students and 107 end of second year students. The second (GEM 8), designed for end of final year students, was in fact piloted on 117 end of third year students, since in that year the fourth year students could not be targeted.<sup>2</sup>

The second phase of the pilot study was carried out in Hong Kong, where 92 ESL first year science students were targeted. The listening summary cloze was in this instance part of the final assessment for the course which consisted of both listening and writing tests.

#### 5.2 Texts

In the first phase of the study carried out in China, the lecturettes were purpose written. The texts were different for the two levels of students and the topics considered were ones which it was felt would be of general interest to the target population. These included ones on nature and wildlife, tourism and places of interest as well as holidays and the way these are celebrated. The main criterion for selection was that the subject matter would be accessible to the students, yet not so familiar that they could complete the summary cloze from their general knowledge.

For GEM 4 the topic selected was *The Hummingbird*. (See Appendix A for the tapescript used and Appendix B for the summary cloze. Details of the other topics are not given as they still form part of a test bank. This text was published in full in Fletcher, 1990.) It was felt that

the students would not know any detail about the bird, but would be able to relate to the subject matter.

In the Hong Kong phase of the study the lecturette used was taken from a radio broadcast. It was from a general science programme which was felt appropriate for the students. It was more authentic than the lecturettes written for China, since the only modifications made were those of editing out some of the detail provided in the talk.

### *5.3 Summaries*

In all cases the summaries were written by one of the testers and then moderated by a group of teachers familiar with the students. The summaries were moderated for comprehensibility as well as to eliminate those deletions which could be completed from the students' general or linguistic knowledge.

The number of deletions varied between tests. For GEM 4 and GEM 8 in China, the summaries contained 20 deletions each, whereas for the Hong Kong test, to facilitate moderation, the number of deletions was increased to 25. A further difference between the tests was the order in which the information was presented in the summary. For GEM 4 the summary followed the same order as the information in the lecturette, whereas for GEM 8 as well as the Hong Kong test the order of the summary did not adhere strictly to that of the lecturette, making the task more challenging. It was felt that by changing the order of information in the summary the task would be more authentic as the students would have to scan their notes to find the relevant information in the same way as they would have to use their notes from a lecture to write an academic essay.

### *5.4 Procedure*

The tests were administered in two stages. During the first stage students were asked to take notes on the lecture they heard. In the instructions they received some guidance as to the particular points they should pay attention to and at the end they were given two minutes to tidy up their notes. They were *not* allowed to see the summary until after the notetaking stage was completed. In the second stage the students

## Listening Summary Cloze Tests

were given the summary and asked to use their notes to complete it. In all cases they were allowed to use a word or phrase to complete the deletion and were informed that they were not required to use the exact words used in the lecturette.

### 6. Results of Piloting

After scoring each of the tests the results were analysed to see how difficult they were for the target population, how the test behaved in comparison to a more traditional listening test and finally how each of the items performed.

The mean scores shown in Table 1 indicate all the tests were of moderate difficulty. A study of comparative difficulty across the tests indicates that GEM 4 was for the year two students of comparable difficulty to the H.K. test for Science students, the mean for the former being 57% while for the latter 56.8%.

**Table 1**  
**Means and Standard Deviations of the Listening**  
**Summary Cloze Tests**

	n	Number of Items	Mean	Std Dev
GIFL				
Yr 1 GEM 4	173	20	8.36	2.76
Yr 2 GEM 4	107	20	11.41	2.85
Yr 3 GEM 8	117	20	9.71	3.67
HKU				
Yr 1	92	25	14.20	3.89

The higher mean score of 11.41 (57%) for year two students as compared to 8.36 (41.8%) for year one students at GEM 4 was to be expected as the test was targeted at year two students. Furthermore, had GEM 8 been administered to end of fourth year students, it is likely their mean scores would also have been higher than the 9.71 (48.5%) achieved by the third years, as they would have completed their fourth and final year of study.

The results of the listening summary cloze were correlated with the other listening subtest administered at the same time. In each case there was a significant correlation: For GEM 4 and GEM 8 they were 0.45 and 0.57 respectively, while for the Hong Kong tests it was 0.48 (all significant at the .001 level), showing that there is some overlap of skills being tested. However, because of the dual nature of the test, it is not surprising that the listening summary cloze test appears to be testing more than listening.

The item analysis carried out on the tests administered in China suggested that only a limited amount of moderation was necessary on both tests. As shown in Table 2, of the 20 items in GEM 4, a maximum of 5 needed to be looked at and possibly changed, whereas for GEM 8 (Table 3) only 2 items were unacceptable.

**Table 2**  
**Summary of Item Analysis for Listening Summary Cloze**

	GEM 4				GEM 8	
	Year 1		Year 2		Year 3	
	Number Items		Number Items		Number Items	
Difficulty	Total	(<.3)*	Total	(<.3)*	Total	(<.3)*
Very Difficult	3	0	2	1	1	0
Difficult	6	0	3	0	5	0
Intermediate	6	1	6	0	10	0
Easy	4	0	5	0	4	1
Very Easy	1	0	4	3	0	0

\*Note: This shows the number of items with a discrimination index of below 0.3, which do not discriminate.

A slightly higher proportion of items, in total 40%, needed moderation after the test was piloted in Hong Kong. This was not seen as disturbing. In fact, it was to be expected on a 25-item subtest that was being used for the first time.

When the item analysis for the listening summary cloze items

## Listening Summary Cloze Tests

(Section 2 of the Hong Kong test) was compared with that for the 15 items that made up Section 1 of the same test, we found that the listening summary cloze, which proved an equally reliable measure, consisted of proportionately more successful items (60%) than Section 1 (27%). It is worth noting here that the larger number of items in Section 2 was a result of these proving easier to write.

### 7. Discussion

The Hong Kong study confirmed the findings of the initial piloting of the test in China, and the results have encouraged both institutions to develop the test further. In Hong Kong the test type is being extended for use with students of other faculties at the University, while in China it has been retained as part of the test battery which will be administered to all students majoring in English throughout the country.

The advantages of the listening summary cloze which have been discussed elsewhere (Lewkowicz, 1991) can be summarized as follows. The test allows for great flexibility in the choice of topic for the talk/lecture and the extent to which the topic is developed. The test can be made easier or more difficult by adjusting the length of the talk, by the selection of topic, and by the means used to develop the arguments presented, which in turn will affect the density of the propositions being discussed. If a radio broadcast is used, as was the case in the Hong Kong pilot study, the text is authentic, though most probably more difficult for the learners to understand—and take notes on—than a purpose written lecture like the ones used in China, since the test setter has no control over the density of information presented.

As implied above, this test format lends itself to a large number of items being written, thus easing pressures on moderation and also allowing for a spread of marks which is very often difficult to achieve in situations where the proficiency level of the students is similar. However, at the moderation stage it is critical that the setter and moderators take into account the possible interdependence of items and the knock-on effect this may have on other items in the summary. Furthermore, if the validity of the measure is not to be compromised,

it is important that setters take due care in ensuring that deletions are made in a way which samples a balance of both lower and higher order listening skills. After all, it must be remembered that this technique is only as good as the deletions made and if these are testing trivial items, then the test will not be a valid one.

Although the marking of the test is largely objective, the learners are not restricted to the words used in the passage: they can use their own words and can use a single word or a phrase to complete any deletion. They need their notes and therefore have a purpose for taking them and in this way the test limits the extent to which learners have to rely on memory to complete the task. Furthermore, in order to check whether the learners are really able to use their notes, the test writer can vary the order of the summary and not follow the information as given in the talk/lecturette.

This test, like any other objective test, is not easy to write. Setters have to be careful in selecting the listening text to make sure it lends itself to a good summary and that there is sufficient redundancy of information to make the task of notetaking possible. Furthermore, they have to ensure that the deletions are appropriate, neither testing general knowledge nor trivial information likely to be missed by the learners. Finally, they have to devise a system whereby the learners cannot fill in the summary completion while listening to the lecturette, otherwise the learners are likely to miss essential information. This can be achieved, for example, by withholding the summary until after the notetaking phase of the test, or preparing the answer booklet in such a way that the learners are prevented from turning to the summary while listening to the text. Whichever method is used, the setters have to ensure clear instructions are given to both the test takers and invigilators. The latter may at the same time wish to draw the learners' attention to the nature of the points to be noted while listening.

Superficially the listening summary cloze may look like a test of fill-in-the-blanks, but the fact that it is a two stage test that requires good notetaking skills and global understanding differentiates it from other

## Listening Summary Cloze Tests

cloze tests. It can have a positive effect on teaching in situations where the objective is to improve students notetaking and listening skills, and it has the added advantage of being a reliable testing instrument.

### 8. Conclusion

The extent to which teachers can try out new testing formats when devising tests depends largely on the nature of the test. If the test is one of major importance to the learners, such as a university entrance examination, then caution is necessary in order to ensure that decisions taken are both valid and reliable.

The quality of any test is dependent on the accuracy of the specification of what the learners need to know and this, as Alderson (1988) has pointed out, is subject to judgements made by those drawing up the specifications. Yet, even with appropriate specifications there is no guarantee that the resultant test will match the qualities of a good test. This is because of the tension that exists between validity and reliability. A test such as the listening summary cloze which is valid in terms of face and content, tests students' ability to listen to spoken discourse and take notes on it, is indirect and hence contaminated. Successful completion of the task is dependent on the students' ability to comprehend the summary utilizing their reading skills, and complete it, using their writing skills. It may therefore prove statistically less reliable than other measures—multiple choice tests, for example—even though it will have a positive washback effect on learning in situations where notetaking is a specified learning task.

Teachers therefore need to be aware of the implications of making compromises on test validity and reliability when designing tests. They also need to realise that in striving towards authenticity of task and a positive washback effect on teaching, decisions have to be taken as to the desired level of test purity. Having taken all these factors into account, they can then design tests that are the most suitable for their learners.

*My thanks go to all the member of the testing team at the Guangzhou*

*Institute of Foreign Languages in 1988/89 who helped develop the test described and made this work possible. My thanks also go to Dr. Desmond Allison and Mr. Peter Falvey for their helpful comments on an earlier draft of this paper.*

**Jo A. Lewkowicz** has been a Senior Language Instructor at the University of Hong Kong since 1989. Prior to that she was a Senior Lecturer at the Guangzhou Institute of Foreign Languages in China. She has taught in many parts of the world including Europe, Africa, and the Middle East.

#### Notes

<sup>1</sup> The proficiency tests were being developed in order to ensure a minimum standard of graduates of English, initially from the Guangzhou Institute of Foreign Languages, but with the view of being used nationwide. The two tests were made up of four papers: Listening & Notetaking, Writing, Speed Reading and Usage which included grammatical accuracy questions as well as reading comprehension.

<sup>2</sup> The tests were ready for administration at the end of the 1988/89 academic year, but were not administered until the beginning of the 1989/90 academic year for political reasons. At the time the Year 3 students were entering their fourth year of study and the Year 4 students had already graduated.

#### References

- Alderson, J. C. (1988). New procedures for validating proficiency tests in ESP? Theory and practice. *Language Testing*, 5, 220-232.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Buck, G. (1988). Testing listening comprehension in Japanese university entrance examinations. *JALT Journal*, 10, 15-42.
- Fletcher, N. (1990, May). Test validation at G.I.F.L.: Towards a more scientific approach. In L. Q. Chen, X. S. Chen & M. Q. Zhang (Eds.). *Proceedings: Second Guangdong Symposium on Teaching English of Foreign Experts/Teachers* (pp. 59-70). Guangzhou.
- Heaton, J. B. (1983). *Writing English language tests* (rev. ed.). London & New York: Longman.
- Lewkowicz, J. A. (1991). Testing listening comprehension: A new approach? *Hong Kong Papers in Linguistics and Language Teaching*, 14, 25-31.
- Rost, M. (1990). *Listening in language learning*. London: Longman.
- Taylor, W. L. (1953) "Cloze procedure": A new tool for measuring readability. *Journalism Quarterly*, 30, 415-433

## Listening Summary Cloze Tests

Weir, C. J. (1988). *Communicative language testing* (Vol. 11, Exeter Linguistic Studies). Exeter: University of Exeter.

### Appendix A

#### Tapescript: The Hummingbird

In this talk I will say a little about hummingbirds, a fascinating family of birds which are found in the Americas. There are in all 319 different species of hummingbird in the world. They are characterized by the fact that they weigh so little, only about one-eighth of an ounce, about the weight of a small coin such as a dime. And as the hummingbirds weigh so little, they are, not surprisingly, also only about 2<sup>1/2</sup> inches long, while the largest of the family is only about 8 inches in length. However, they all have an exceptionally long bill or beak which the bird uses to sip nectar from flowers or to catch insects.

An extraordinary fact about the hummingbird is that their energy output is the greatest per unit of weight of any warm-blooded animal. Whereas a man weighing 170 lbs produces 3500 calories of energy per day, the hummingbird produces 155,000 calories per day, that is nearly 45 times as much. In order to fulfil this tremendous need for energy, the hummingbird must consume half its weight in sugar daily. A typical day's supply of food would come from 1000 fuchsia or hibiscus blossoms, so much of the bird's day is spent on feeding. Incidentally, hummingbirds are especially attracted to red. They will buzz at a woman's mouth if she is wearing red lipstick!

Another noticeable characteristic of the hummingbird is that they have larger flying muscles in proportion to their weight than those of any other bird. The movement of their wings is extremely fast - up to 200 beats a second during courtship, 75 beats per second during normal flights when collecting nectar or blossoms. The rapid beating of the wings creates a distinct humming sound which is why the bird is called a hummingbird.

Of the 319 species of hummingbird, most are found in South America but a dozen species live in the U.S.A. Of these 12, only one lives east of the Mississippi River and nearly half favour Southern California. The greatest traveller among these birds is the Rufous hummingbird, spelt R O U F U S, the Rufous hummingbird. It migrates at least 2000 miles from Alaska to Mexico and it is believed that this ruby-throated hummingbird flies non-stop across the Gulf of Mexico, a distance of 500 miles.

Hummingbirds, especially the males, are magnificently coloured, varying from deep metallic greens which tend to be the dominant colour, to intense reds and yellows. Their fairylike beauty is impressive especially in light of their physical capacities. They are sturdy and quick. They are also the only bird that can fly backwards as well as forwards. They average a ten-year life span. And they can live both in the heat of the Mexican desert and the cold of the Andes

Mountains.

So as you can see, it is not surprising that these birds have caught the attention of man and have been the subject of much study and concern.

Let me now summarize the main facts again. There are 319 species of hummingbird in the world and most of them live in South America but a few are to be found in North America. They are extremely small birds, ranging from 2<sup>1</sup>/<sub>2</sub> to 8 inches in length. Yet for their size they have a very long beak which they use to collect nectar and catch insects. These birds have strong flying muscles which allow them to fly long distances. And it is the humming noise they make while moving their wings that has gained them their name. Hummingbirds, especially the males, are magnificently coloured and one feature that distinguishes them from other birds is their ability to fly backwards.

Appendix B

Summary Cloze Passage

Using the notes you have made during the lecture, complete the summary below. You may use either a word or phrase to fill in each of the blanks. You do not need to use the exact words that were used in the passage.

There are (1)\_\_\_\_\_ different species of hummingbird in the world of which twelve live in (2)\_\_\_\_\_. They derive their name from the sound they make by (3)\_\_\_\_\_.

These birds, despite their small size, have (4)\_\_\_\_\_ which they use for collecting nectar or (5)\_\_\_\_\_. In order to produce the energy they require, the birds spend much of their day feeding and it is not uncommon to see them (6)\_\_\_\_\_ above blossoms, especially those like the hibiscus that are (7)\_\_\_\_\_ in colour.

The hummingbirds' daily energy output is approximately 155,000 calories, which is (8)\_\_\_\_\_ than that of the average size man. To obtain this, they must consume (9)\_\_\_\_\_ their own body weight in sugar daily, that is about one-sixteenth of an ounce per day.

None of the hummingbird family is large; they range from (10)\_\_\_\_\_ to (11)\_\_\_\_\_ inches in length, yet they are well-built with (12)\_\_\_\_\_ which enable them to travel relatively long distances. The (13)\_\_\_\_\_ hummingbird, for example, migrates 2,000 miles from Alaska to Mexico crossing the (14)\_\_\_\_\_ stretch of the Gulf of Mexico nonstop.

One noticeable feature of the hummingbird is its magnificent colouring. The birds vary from (15)\_\_\_\_\_ which is the dominant colour, to (16)\_\_\_\_\_. As with other birds, the males tend to be more (17)\_\_\_\_\_ than the females. But unlike any other bird the hummingbird can (18)\_\_\_\_\_.

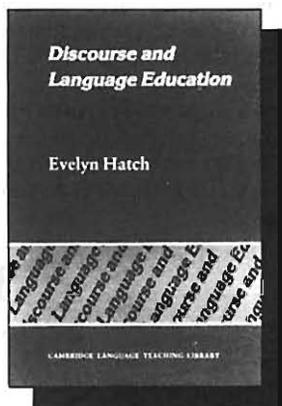
Two further factors worth noting are that the hummingbird has an average lifespan of ten years, and its habitat may be very diverse from (19)\_\_\_\_\_ to the cold of (20)\_\_\_\_\_.

## Discourse and Language Education

Evelyn Hatch

*Discourse and Language Education* gives practical experience in analysing discourse. The book includes:

- analyses of *spoken* language and the structure of 'scripts'
- the study of *written* language – the rhetorical structure of compositions or more informal written material such as personal letters.

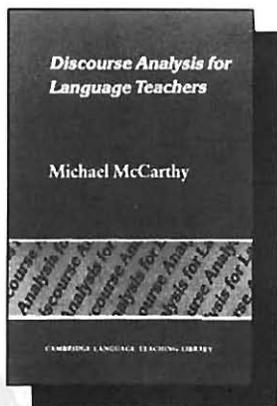


## Discourse Analysis for Language Teachers

Michael McCarthy

*Discourse Analysis for Language Teachers* gives a practical introduction to the field of discourse analysis and its relevance for language teaching. The book:

- examines how discourse analysts approach spoken and written language
- outlines and evaluates different models of analysis
- includes chapters on discourse-oriented approaches to grammar, vocabulary and phonology



The text is interspersed with reader activities with guidance on appropriate responses at the end of the book.

For further information, please contact:  
David Fisher, Cambridge University Press,  
United Publishers Services Ltd, Kenkyu-sha  
Building, 9 Kanda Surugadai 2-chome,  
Chiyoda-ku, Tokyo 101.  
TELEPHONE: (03) 3295-5875

**ACTIONARY!**, the fun dictionary has combined the most tedious task of memorizing vocabulary with the visual media. It is a 6 volume set, 30 minutes each, with a basic theme running through each volume, levelling up in difficulty with each volume. Dynamics are emphasized through the use of location film shot in Seattle, studio and animation as well as segment lengths, rhythm and a variety of music.

- Phonics practice especially with vowels and vowel combinations.
- Vol.4,5,6 include a Quiz, which allows the viewer to participate.
- Many of the words are spoken in a situation so that the student can have a better grasp of the meaning of the word and how they can be used.



written and directed by  
**YOKO NARAHASHI**  
(MLS)



THE VIDEO DICTIONARY

# ACTIONARY!

- 507 compulsory words required for Junior High School by the Educational Board.
- A text accompanies each volume which follows the order of words shown.

- ① My Day
- ② Friends
- ③ City
- ④ Time and Places
- ⑤ Feelings
- ⑥ Activities



¥ 70,040.- (including tax)

VIDEO (HiFi stereo: color 30 min. each) 6 vols. + TEXT 7 vols.

263-1 Tsuruma, Machida-shi, Tokyo 194 WALKERS COMPANY INC. TEL.(0427)99-3393  
2-3-6 Ginza, Chuo-ku, Tokyo 104 OKURA & CO., LTD. / E.D.S. OKURA TEL.(03)3564-6197

# Using Learner Evaluations to Improve Language Teaching

Keith Maurice

*Florida State University*

This article looks at teacher evaluation and program development from a learner-oriented perspective. After briefly reviewing the literature related to observation as an element in teacher evaluation and development, the article explores and then advocates the use of learners as course and teacher evaluators as a central component in language programs. It discusses the rationale for, research on, problems with, and ways of changing and combining three different roles: learners as evaluators, evaluators and supervisors as teachers, and teachers as learners. Mini-cases are used to illustrate the main issues involved.

## 学習者の評価を使った語学教育の改善

本稿は学習者の観点から教師評価と教授内容の進展を考察する。まず学習者の評価を教師評価と教育課程の一要素として検討する。

次に学習者がコースや教師の評価者となり、語学学習の中で中心的要素となる可能性を探り盟道する。その主張の論理的根拠、研究、問題点、方法を論じ、かつ評価者も教師も自ら変化すると同時に両者の役割を合わせ持つ必要性を考察する。又、実例をあげて関連した主要問題点を説明する。

## 1. Introduction

Improving the effectiveness of instruction in language programs can be very difficult for a variety of reasons. To become more effective implies an understanding of what is important, an ability to evaluate what is important, and then the skill to change whatever needs changing. None of this is easy. As Allwright and Bailey (1991) put it:

Becoming more effective cannot be a simple linear matter. We cannot expect to reach more understanding one day, and then

simply be more effective the next. We have to work continually for increased understanding, and work simultaneously for enhanced effectiveness, in a constant cycle with no starting point and certainly no single and triumphant finishing point. (p. 196)

What can we do as teachers—and as supervisors who evaluate teaching—to improve both instructional and program effectiveness on a continuing basis? This paper looks briefly at various ideas that have been proposed, chiefly dealing with observation, and then investigates the use of learner evaluations of teaching as one component that can offer important information to teachers and evaluators alike in their quest to understand learner perceptions and then to work on improving instructional services.

The focus of the paper is on the evaluation of teachers. McLaughlin and Pfeifer (1988) state that “Teacher evaluation is . . . pursued as a potent strategy for enhancing both quality and control of . . . education,” but “teachers share neither this enthusiasm nor these expectations” (p. 1). McNeil (1981) goes further to say that teachers fear evaluation. Stiggins and Duke (1988) sew the threads together by saying that “The paradox of teacher evaluation is that it holds the potential to help nearly every teacher improve, yet in actual practice it helps almost no one” (p. 1). Clearly, then, this is an area that deserves more attention by teachers and evaluators alike.

Teacher evaluation has traditionally been done, at least in part, by observations of teachers in the classroom. Sheal (1989) notes several problems related to this arrangement:

1. Most classroom observations are conducted by administrators rather than by practicing teachers.
2. Much of the observation that goes on is unsystematic and subjective.
3. Most observation is for teacher-evaluation purposes, with the result that teachers generally regard observation as a threat.
4. Post-observation meetings tend to focus on the teacher's

## Learner Evaluations and Language Teaching

behavior rather than on developing the teacher's skills. As feedback from observers is often subjective, impressionist, and evaluative, teachers tend to react in defensive ways, and given this atmosphere, even useful feedback is often "not heard." (p. 93)

In various ways, then, using observations for traditional evaluation purposes limits their usefulness for evaluation, while ignoring the pedagogical value of classroom observation. Several writers (see Fanselow, 1987, p. 20; Sheal, 1989, p. 92; Williams, 1989, p. 85) have called for descriptive and developmental approaches to classroom observation which downplay or eliminate the evaluative function altogether. Allwright and Bailey (1991, p. 198) see classroom research done by the teachers themselves—"exploratory teaching"—as a solid way to keep improving. Richards (1990, p. 119), calling it "self-monitoring," also sees benefits in this approach, noting that it encourages the use of reflection about one's teaching, helps teachers become more realistic in their views of their own teaching, and makes the teacher, and not the evaluator, responsible for his or her own improvement.

Changing the focus of classroom observations to exploration and self-monitoring is a positive direction for the language teaching profession because it encourages teacher development. However, along with teacher development, there is also a need for program development. Administrators still need information about teaching and teachers in order to make fundamental program and personnel decisions. For example, if one teacher is having major problems in getting along with students in course after course, the administrator may feel obliged to take action in one way or another for the benefit of the program and the learners. It may take years for an ineffective teacher to become an effective one, but administrators may have to deal with the consequences of ineffective teaching on a recurring basis. As such, using learner evaluations of teaching is one component of an evaluation system that can aid administrators in making program decisions and also assist teachers in becoming more aware of certain matters that are

key to their own development, such as how the learners perceive their instruction.

The central thrust of the paper is for the need of teachers, evaluators, and learners to change and combine roles in order to enhance what goes on in the classroom. There is much we can learn from the learners if we step out of our own roles long enough to reflect on what they have to say. This shifting of roles is seen as essential but not at all common. As Sheal (1989) notes, "the world of education tends to be a 'world unto itself' and teachers and their supervisors often fail to realize that they may learn something from other worlds" (p. 102). It appears easy, for example, for educators to assume that the roles of learner, teacher, and evaluator are completely distinct and different from one another. Though each role does carry with it certain explicit and implicit tasks and responsibilities, there is also much that members of each group can learn from the others.

The paper is divided into three sections: learners as evaluators, evaluators as teachers, and teachers as learners. The context involves both university and adult language learners. As such, the assumptions are more androgogical (human oriented) than pedagogical. This distinction is relevant because of the importance of learner maturity to the notion of learners evaluating teachers. The mini-cases described involve English as the target language, mostly as a foreign language, and ESP teacher education. The implications, however, fit other foreign and second language teaching situations as well.

## **2. Learners as Evaluators**

### *2.1 Rationale for Using Learners as Evaluators*

In university or workplace ESL/EFL settings, learners ostensibly have reasons for and choices in making decisions about learning another language. Even where the learners are required to study English, they can still generally be assumed to be reasonably mature, fully-functioning individuals capable of engaging in meaningful discourse about their language learning process. As such, it seems reasonable to include their views about the course in any evaluation; in fact, this is being done in many educational settings (Pennington & Young, 1989).

## Learner Evaluations and Language Teaching

According to Aleamoni (1981), there are several key reasons to get information from the learners, including (emphasis added):

1. *Students are the main source of information* about (a) accomplishment of important educational goals, such as the development of motivation for continued learning, and (b) areas of rapport, degrees of communication, and the existence of problems between instructors and students.
2. *Students are the most logical evaluators of the quality and effectiveness of and satisfaction with those course elements* (instructor, textbook, homework, course content, instruction, student interest, . . . attention, and general attitude).
3. *Student ratings provide a means of communicating between students and instructor*, which in large institutions may not exist in other forms.
4. *Student demands for information* about instructors and courses . . . *encourage instructional improvement.* (p. 111)

In discussing the learners' role in change stemming from evaluation, Loew (1979) notes the relation between involving learners at the beginning of the process and their subsequent commitment to the results of that process. For some educators, this may entail a substantial change in perspective—from seeing learners as subordinates to seeing them as partners or clients. That change, however, can yield benefits to all.

### *2.2 Concerns and Research Findings about Learners as Evaluators*

Having stated the basic reasons for involving the learners in the evaluation process, it becomes important to look at the possible problems related to using learners in this way. Aleamoni (1981, p. 111) addresses concerns that have been expressed by teachers in various fields. The concerns are listed below, along with Aleamoni's summary of the research findings in italics:

1. Learner judgments are inconsistent due to “immaturity, lack of experience, and capriciousness.” *False—Judgments are very consistent.*
2. Only those who are published researchers can properly teach or evaluate other teachers. *No clear trend is evident.*
3. Learners use popularity rather than effectiveness as their criterion. *False—Learners are precise evaluators.*
4. Learners need to be away from a course to gain the perspective needed to evaluate accurately. *It is difficult to get solid information here, but what research there is indicates agreement between present and past students.*
5. The rating forms lack reliability and validity. *When the forms have been developed professionally, the reliability is high. Most evidence suggests validity as well.*
6. Extraneous variables such as class size, the sex of learner and teacher, time of the class, course level, teacher’s rank, course as requirement or elective, and learner as major or nonmajor will affect results. *Research into size of class, sex, course as requirement, course level, and teacher rank is mixed, while that into time and student major indicates basically no effect.*
7. Learners’ grades correlate highly with ratings given. *This varies widely, but is most often weak.*
8. Midterm learner ratings are not useful in improving teaching. *Research on this is mixed.*

The concerns listed above certainly point to teacher suspicions about having learners evaluate them. However, the research indicates that some of the concerns may be unfounded while others are much more complicated than some teachers think.

In the field of second/foreign language teaching, cultural differences and other matters may sometimes come into the evaluation picture. Sharp (1990) notes five problems relevant to an EFL course in Brunei Darussalam:

## Learner Evaluations and Language Teaching

1. Asking a student about your course may suggest that you do not know what you are doing.
2. Students may not give an honest reply to questions; they may respond in the way they think you want them to respond.
3. Students may display some personal antagonism towards individual teachers or situations.
4. Students may feel that commenting on the effectiveness of the course is not their concern.
5. Students may even feel that adverse comments may have some future negative effect on their grades. (p. 135)

The issues of foreign language learner expectations, fears, and diplomatic maneuvering are well worth looking into and may be very much culturally, institutionally, or even personally defined. Thus, using learners' views as evaluative input may be very complicated in some contexts in some parts of the world. In such contexts, much preparation and persuasion on the part of the administrator involved may be needed to begin the process. Nevertheless, the rationale for and the research on using learner evaluations in university and adult settings are generally too strong to be dismissed lightly. The author's own experiences with Japanese, Thai, and learners from many other nations mirror closely what Aleamoni (1981) has noted.

### *2.3 Ways of Changing the Learner's Role*

With all the perceived problems noted, it might be easy to dismiss the use of learners as evaluators as a wonderful idea which is not realistic. That would be analogous to throwing the baby out with the bathwater, however, because some problems are more imagined than real, and others can be dealt with to negate any major drawbacks that may exist. Sharp (1990), for example, mentioned that talking over the purposes of his questionnaires with the learners and emphasizing the use of anonymity helped him to avoid some of the problems.

Waters (1987) and Lewkowicz and Moon (1985) have noted specific ways of involving learners in the evaluation process. Lewkowicz and

Moon (1985) state that in the communicative approach “the learners can share some of the management responsibilities with the teacher” (p. 47). This seems to be both a reasonable notion for the adult learning situation *and* a natural cause of concern for individual teachers worried about having even more pressures put upon them. It could also be said that the whole move toward communicative teaching has involved a paradigm shift so huge for some teachers that the idea of sharing classroom *and* evaluation responsibilities would be seen as threatening. Nevertheless, *if* the focus is on the learning process, then it seems essential to involve the learner *and* to work towards enabling the learner to evaluate various aspects of the learning process (Breen and Candlin, 1980).

Some of the ways of using learners as evaluators include (a) using anonymous end-of-course rating scales; (b) using anonymous end-of-course open-ended questionnaires; (c) having informal discussions with learners throughout the course (Waters, 1987); (d) having student government committees; (e) having student-teacher curriculum committees (Loew, 1979); (f) having learners write dialogue journals throughout the course; (g) interviewing learners at the end of the course or at the end of their stay in the program; and (h) giving all learners questionnaires at the beginning of the course so that even dropouts may provide feedback (Scriven, 1981). There are pros and cons, and cautions, with each way mentioned.

Rating scales are probably the most common way of getting learners to evaluate teachers and courses. Doyle (1983) lists the pros and cons as:

PROS	CONS
1. Easily analyzed & stored	1. Unable to reflect subtle thoughts
2. Efficient (quick & cheap)	2. Do not promote reflective responses
3. Structure allows precise & uniform data-gathering	3. Can be seen as tedious, repetitive & irrelevant
4. Detailed guidelines available	4. Do not facilitate dialogue

## Learner Evaluations and Language Teaching

The cons noted above are important and can be neutralized by other more responsive measures, such as informal discussions, interviews, and quick action on problem areas mentioned. As in other types of evaluation, triangulating the data is valuable here no matter which technique is the main focus of the effort.

Waters (1987) mentions definite benefits of involving learners in course evaluation throughout the course, through the use of regular, informal feedback sessions. Such benefits include opportunities to respond quickly to any weak points in the course, the establishment of trust, and building commitment to the course goals into the process. As Hutchinson and Waters (1987) note, “an open and trusting relationship between the teacher and the students is the best basis for promoting frank and useful feedback” (p. 154). Thus, student involvement can include both reaction to the teaching *and* participation in the planning of what goes on in the classroom.

### *2.4 Uses of Learner Evaluations*

Learner evaluations of course elements such as content, materials, teachers, and teaching can be used both for formative and summative purposes. For formative purposes, such evaluations can be very helpful in adjusting the courses to make them more appropriate to the learners. Such adjustments can be made in teaching techniques, methodology, materials, and in matching content to learner needs and interests. Teachers, curriculum coordinators, and administrators can become blinded to certain practical issues that the learners see very well. As such, getting accurate information from the learners about their perceptions of what is going on can be of great benefit to the faculty and staff—*if* the evaluations are interpreted and used responsibly.

In working towards the fair and worthwhile use of student ratings, Pennington and Young (1989) advise that the evaluation instruments and procedures be made by evaluators familiar with language teaching, that the ratings include room for open-ended responses, and that learners be oriented to both the purposes and the content of the evaluation. They also note that:

Student evaluations of (language) teachers can thus best serve formative purposes if employed in a context of strong administrative support for faculty members, public rewards for outstanding performance, and opportunities for training to achieve growth as a language teaching professional (p. 629)

It remains very important to remember that student ratings are not seen as the sole basis for teacher evaluation, but as one very powerful and relatively direct way of viewing the process, especially for formative purposes. For summative purposes, triangulating the data through the use of other measures—such as teacher interviews, classroom observation, peer review, self-evaluation—is the key to evaluating both fairly and effectively.

### *2.5 Cases of Learners as Evaluators*

Real cases from EFL and ESP teacher education will be discussed briefly to illustrate both the potential and the problems of the evaluation process noted in this paper. In this section, only the learner as evaluator will be discussed; later, the cases will be expanded to include the other roles.

#### *2.5.1 The ineffective teacher with little support*

In a language teaching company in Japan in which classes were conducted at client corporations, one teacher consistently had classes with high dropout rates. Though no formal student evaluations were used by the company, student-generated complaints about the teacher were frequently reported by the client representatives. These complaints described a teacher who was very nervous and lacking in confidence. This resulted in a tendency to mumble and fidget frequently. In addition, the teacher lacked focus in the classroom. He did not indicate to students the purpose of the various activities; he also did not plan lessons or organize the classroom time well. He often rushed through the regular lesson and then fumbled around for topics to discuss or activities to do. This resulted in even more mumbling and fidgeting.

Discussions with the client representatives helped the teacher's supervisor to assist the teacher by giving him much more guidance in

## Learner Evaluations and Language Teaching

overcoming fear, speaking clearly, planning lessons, organizing time effectively, and using techniques to get the students more involved in the classroom. In addition to giving the teacher guidance, the supervisor began creating supplementary materials for possible use by all teachers in the program. These materials provided the students with more focus for their language activities, gave the nervous teacher a “crutch” to rely on, caused the beginning of a practical resource file for all teachers, and indicated to the client company’s representatives that the educators took client concerns seriously.

### *2.5.2 The course that fit the program’s objectives but which generated unsatisfactory process and results.*

In an ESP teacher training program, one course was designed to familiarize English teachers with the world of science. This course specifically fit the particular program’s objectives in preparing teachers to enter university-oriented teaching. The course was handled by an ESP teacher who had scientists talk to the learners about various scientific disciplines. Student ratings were used in the program and, according to the teacher, had not indicated any problems with the course. However, a program evaluator’s interviews with former students of the course revealed intense and antagonistic feelings about the course. The majority of the students indicated a desire, often with much emotion, to eliminate the course from the program. Although the specific reasons for this desire were not obtained, the information from the interviews was valuable to the evaluator in several ways. First, the ratings previously used were not anonymous; the students were instructed to sign their names on their forms, violating one of the basic tenets of evaluation. Second, the complete mismatch between learner and teacher statements about the course clearly indicated a need to rethink the rationale and methodology used in the course. Third, the unexpected discovery by the evaluator of problems with the course helped to broaden the evaluator’s own perspective in conducting other evaluative and administrative work.

### 3. Evaluators as Teachers

Once the information from students is obtained, the supervisor or evaluator analyzes the information and then tries to use it in ways that can improve the program. The evaluator as teacher is an essential part of this effort. As Cronbach (1983) notes, "Teaching begins when the evaluator first sits down with members of the policy-shaping community to elicit their questions. It continues during every contact the evaluator has with program participants or with others in his audience" (p. 9). The commonly perceived role of evaluator as judge, while important, can serve to so intimidate or antagonize teachers unnecessarily that it can get in the way of working effectively if it is not combined with the supportive work of teaching the teachers. Too much emphasis on the judge role can leave the evaluator with the image of a Darth Vader, a dark and sinister force to be resisted; such an image serves only to complicate the evaluation process and create problems for interpreting and using the information obtained.

#### *3.1 Rationale for Supervisors/Evaluators Being Teachers*

There are at least four related reasons why evaluators need to function as teachers. First, as mentioned above, the evaluator needs to reduce fears and suspicions so that the evaluator can do his or her job without every move being closely scrutinized by teachers (McNeil, 1981). Second is the need to inform the teachers of the goals, criteria, procedures, and conflicting desires that define and constrain the evaluator (Darling-Hammond, 1983). Third is the need to establish and encourage trust and open communication with the teachers involved (McLaughlin & Pfeifer, 1988). Fourth is the need to persuade the various stakeholders, including teachers, to take action on the evaluation once the recommendations have been given (McLaughlin and Pfeifer, 1988). Any evaluator who ignores these needs risks losing credibility and seeing recommendations or directives ignored.

Supervisors, by the very nature of their work, evaluate in order to enhance program development. Dealing with teachers who are perceived as being ineffective can be at times a very difficult and unpleas-

## Learner Evaluations and Language Teaching

ant task, but whatever the supervisor does has implications for the program and for virtually everyone involved in it. Tolerating poor performance and sidestepping the problems sends a clear signal to teachers and learners; it teaches them that learner evaluations have no tangible relevance to or influence on their day-to-day lives. At the other end of the spectrum, reassigning an ineffective teacher or removing that teacher from the program entirely sends a signal that ongoing evaluations by learners are extremely important elements of the learning/teaching environment. In the middle, of course, lies the path of working with the affected teacher to improve performance as perceived by the learners. Ongoing teacher development is at the center of the evaluator's task.

### *3.2 Characteristics of Supervisors/Evaluators as Teachers*

What characteristics does an evaluator need to teach teachers? Brock (1981) notes that the most important relate to “a commitment to student learning, an abiding curiosity about the relationship between teacher, student, and subject matter, an empathic disposition, a knowledge of local resources, a tendency toward self-disclosure, and effective interpersonal skills” (p. 239). Though these are very important, a sense of balance also seems important. For example, some educators indulge in so much self-disclosure that the purpose of the interaction may be lost in the process. Balance is needed to protect against the polar extremes of too little or too much self-disclosure.

If evaluators need to be teachers, how should they teach? McLaughlin and Pfeifer (1988) talk about the need for evaluators to enable others to go through “the process of unfreezing, of reexamining the understandings, beliefs, and practices fundamental to the institution [and note that] some kind of triggering event appears necessary” (p. 30). The concept of unfreezing is one that many educators are already familiar with, at least intuitively. It is both an appealing and powerful analogy. Helping others to unfreeze, however, is not an easy task. Any evaluator who has had to face teachers who have not succeeded in earning the appreciation or respect of the learners knows the emotional turmoil that can be caused by negative evaluations. As unpleasant as this can

sometimes be, evaluations by learners and discussions afterward between teachers and supervisors can serve as triggering events needed to begin the process of development.

Cronbach (1983) adds that “educating is as much a matter of raising questions as of providing answers. Especially where the topic is value-laden, the educator’s responsibility is to help others ask better questions and determine what actions are appropriate to their own aims” (p. 9). He goes on to say that the evaluator should “reduce uncertainties, but . . . should also challenge simplistic views” (p. 10). One way the supervisor can do this is to analyze and compare learner evaluations from many classes in depth, including such matters as time of the class, skill area taught, and how the same learners rated different teachers. What is learned from this can then help in responding to defensive statements and clarifying various matters to teachers. For example, an afternoon teacher may get unsatisfactory ratings from the learners and then say that the reason for the bad ratings is the time of the class. By analyzing and comparing data from the evaluations beforehand, the supervisor will be ready to respond to the statement and hopefully trigger, or re-trigger, the process of unfreezing.

### *3.3 Problems with Supervisors/Evaluators as Teachers*

Since both evaluation and teaching are intensely human endeavors, the whole range of human emotions, weaknesses, and pitfalls serves to complicate the situation of the evaluator teaching the teacher. Brock (1981) mentions some of the problems facing the evaluator as helper, such as (a) irrelevance of the advice; (b) lack of awareness of the other person’s resistance to the “help”; (c) mixed signals due to rapport building and reluctance to criticize; (d) the other person’s desire to keep things hidden; and (e) the other teacher’s tendency to deny or shift blame.

Irrelevant advice can sometimes be heard from evaluators who use too much self-disclosure in their interactions. The message may be an irrelevant, anecdotal story rather than a persuasive recommendation of ways to improve teaching. Self-disclosure may also create mixed

## Learner Evaluations and Language Teaching

signals. Both the rapport-building process and natural reluctance of some evaluators to criticize can work against getting the real message across to the teacher. Bridges (1986) notes that assurances of job security, ambiguity in teacher evaluation, and the desire to avoid conflict and discomfort can confuse the specific teacher and demoralize other teachers, who may feel that incompetence is being rewarded while their efforts and results are being ignored.

The second, fourth, and fifth problems mentioned above relate to the difficulty of actually getting ineffective teachers to improve. Bridges (1986) is critical of “salvage attempts” to improve the performance of poor teachers. He notes that such attempts are characterized by “unmuted criticism, defensive reaction, behavior specification, limited assistance, restrained support, extensive documentation, and little improvement” (p. 48)—in other words, much ado about nothing. Effective teaching by evaluators, then, is complex. The factors which affect the evaluator as teacher, combined with those which affect the teacher as learner, make the situation simultaneously extremely important and extremely delicate.

### *3.4 Key Points For Evaluators as Teachers to Remember*

It is not only what is communicated but how, when, where, and for what ends. McLaughlin and Pfeifer (1988) note the key elements of timeliness, specificity, credibility, and intent. The evaluation should include feedback to the teacher soon after the evaluation, deal with specific concerns, come from someone the teacher respects, and be given in ways the teacher feels are supportive. They also note that:

Teacher evaluation strategies, through open communication, interaction, and discussion, provide the kinds of professional stimulation and feedback that support individual growth consistent with institutional goals and values. Organizational control achieved through such normative means is the more enduring, robust, and predictable. (p. 84)

Loew (1979) discusses various teacher characteristics that need to be addressed in the process of modifying programs after evaluation. The

evaluator as teacher needs to address them as well. They include skills and knowledge needed for effective change, the process of acquisition of these skills, the encouragement of teacher creativity, and the process of feedback as changes are made. Cronbach (1983) also notes the value of communication skills and planning for communication with teachers. He provides a useful list of questions for evaluators:

- Did each fraction of the audience attend to the message?
- Did each understand it?
- Did each find it credible?
- Were the significant questions answered as well as possible?
- Did the answers alter the preconceptions of the audience?
- Was the dialogue leading to the decisions enriched and elevated as a consequence of the evaluation? (p. 11)

In all of the comments above, the evaluator is seen as a colleague or as a helper, not as a jealous critic or judge from on high. To the extent that the evaluator is respected and is perceived as a positive person, the message is more likely to be taken positively.

### *3.5 Cases of Supervisors/Evaluators as Teachers*

Three cases will be discussed here, including the two mentioned previously.

#### *3.5.1 The ineffective teacher with little support*

In this case, the curriculum coordinator served as the supervisor in charge of dealing with the situation. Both the supervisor and the teacher were Americans. Private discussions were held with the ineffective teacher about the concerns and complaints of the clients. Since general rapport between the two was good, communication during the discussion was open, friendly, and frank. Upon hearing the concerns, the teacher immediately expressed regret and vowed to try to do better. The supervisor showed concern for the teacher's well-being and offered some general ideas and specific techniques and activities to produce better results in the classroom. The supervisor also encouraged the teacher to create worksheets to use with the book and volunteered to create others.

## Learner Evaluations and Language Teaching

In the ensuing weeks, repeated discussions were held with the teacher to help with lesson planning and to discuss various ways of dealing with classroom situations. In addition, many new worksheets for learners were developed, mostly by the supervisor, both to help the particular teacher and to begin a teachers' resource file for all teachers in that branch of the company.

The supervisor used this case of an ineffective teacher to generate new processes and improvements for others. In so doing, more teaching was needed to persuade other teachers to join in the effort to create a teachers' resource file. The philosophy underlying the effort was "All for one and one for all," but some teachers resisted the idea, including a couple who actively resented any calls for sharing their good ideas with others. However, after a few months, when the resource file was filling with supplementary text exercises, vocabulary and grammar worksheets, language games and quizzes, and descriptions of teaching techniques and activities, the resistance melted away. Teachers recognized the value of the file to their own teaching.

The supervisor in this case spoke to the teacher in a timely manner, discussed very specific concerns and ways of dealing with those concerns, demonstrated his credibility through general knowledge and ability to help, and displayed good intentions by offering positive suggestions and then helping to implement some of the suggestions. As such, McLaughlin and Pfeifer's (1988) four elements were covered. Using Cronbach's (1983) questions to look at the case, it can be seen that the supervisor's communication with the ineffective teacher was solid. With some of the other teachers, however, the message was not understood completely in the beginning, nor seen as credible. The effort to teach them took more time to be successful.

### *3.5.2 The course that fit the program's objectives but which generated unsatisfactory process and results*

In this case, the evaluator arranged a private meeting with the teacher to discuss the very negative findings. These two people were from different cultural backgrounds. The meeting was held prior to the

evaluation report being turned in to the department head and becoming public within the department. The evaluator discussed his own experiences and past problems in coordinating other speakers and empathized with the teacher's predicament in this regard. The evaluator also probed the teacher for possible factors that could explain the intensity of the student views. Finally, the evaluator and teacher discussed their thoughts on modifying the course. At the meeting, the communication seemed open and productive. The teacher asked for the evaluator to help teach the course in the future.

### *3.5.3 The ineffective teacher who did not admit the problem*

One university teacher repeatedly received poor ratings from students in every class she taught. At the time a new evaluator first talked to her about the situation, her problems were well-known throughout the language teaching department. After documenting the poor student ratings for one semester, the evaluator talked with her about the ratings and tried to empathize with her situation. The teacher, however, was already defensive and stated that it was the students' problem, not hers. The evaluator tried to explain that it was a problem not only for the students, but also the teacher and the evaluator/supervisor. This discussion was intended as a gentle yet straight-forward awareness building session.

During the next semester, several influential teachers in the department advised the supervisor to dismiss the teacher from that particular program. The supervisor listened quietly and thought carefully about what was best for the students, for the teacher, for teacher development generally, and for the department. The effect of any decision on future departmental politics had to be considered as well. After poor student ratings came in again that semester, the evaluator again went to talk with the teacher about the problem. Again, the teacher was defensive. The evaluator tried to find a solution that would help the teacher improve her teaching while saving face and also help the program. He suggested and she agreed to have another teacher become co-teacher with her in the course, with the idea of the two teachers working together. This case illustrates both the process of the evaluator as teacher *and* the problems involved in dealing with defensive teachers in politicized situations.

# Learner Evaluations and Language Teaching

## 4. Teachers as Learners

Teachers generally fear or are suspicious of evaluations and often see them as punitive rather than instructive. Omaggio, et al. (1979) note that some teachers view evaluation as an “invasion of privacy” (p. 236). From an organizational viewpoint, this may seem petty and short-sighted; however, from a teacher’s viewpoint, the fears and suspicions are often justified. Evaluations from supervisory observers are often just snapshots and may not be very good snapshots at that. Administrative biases and blindspots may work against the program goals rather than for them, leaving some teachers feeling like Don Quixote, doing their best against impossible odds.

Nevertheless, *if* we can obtain reasonably valid and reliable evaluations from learners, and *if* the evaluators can show that they want to improve the program and the teaching that goes on within it, *then* it seems obvious that the teacher should be willing to learn and keep on learning throughout his or her teaching career. Indeed, even if the first two conditions are not met, it behooves every teacher to keep learning whenever and however for self-actualization purposes. As Brock (1981) notes, “Motivation is probably a key characteristic. Those who naturally seek to understand their effects on students and who are committed to finding improved methods of teaching are likely to benefit from evaluation-based development” (p. 240).

### 4.1 Rationale for Teachers Becoming Learners

Performance in education is not constant over time and the process of learning is, or should be, lifelong. Reflecting the need for teachers to keep learning throughout their professional careers, Smith (1977) notes four specific reasons for doing so, noting that learning can:

1. remedy the teacher’s deficiencies arising out of defects in . . . initial teacher training preparation,
2. advance the teacher’s skills and pedagogical knowledge required for new teaching roles,
3. advance and update the teacher’s knowledge of subject matter,

4. train the teacher for non-tutorial positions. (pp. 179-180)

Given the state of affairs in much of teacher education, Smith goes on to note that lifelong learning often needs to be remedial in nature.

In discussing teacher education, Widdowson and Candlin (1990) discuss the terms awareness, knowledge, understanding, concern, and action. If we think of ourselves as teachers, we can also ask which of the above qualities we have. To improve teaching over time, all of the five qualities are needed and none can be taken for granted. Knowledge does not necessarily lead to concern, for example, and concern does not necessarily lead to positive or successful action.

Awareness, however, is one of the keys to this whole process. Teachers need to be aware of a variety of personal characteristics and course elements and processes as they do their work. The teacher as learner needs to be able to use self-assessment in a reflective way (Loew, 1979). As Murphy (1985) states, "A teacher who has awareness understands the processes, as far as that is possible with our limited knowledge. This teacher realizes that we do not know all the answers and cannot be answerable for everything, that defensiveness and secrecy do not help" (p. 14). These qualities describe the professional perspective, but insecurities and suspiciousness abound in many ESL/EFL contexts. However, if teachers can overcome their fears and take some risks, the benefits can be substantial in both personal/professional growth *and* organizational effectiveness (Nunan, 1988; Stenhouse, 1975). As Nunan (1988) notes, "it is not enough that teachers' work be studied, they need to study it themselves" (p. 147).

#### *4.2. Methods of Teachers Learning*

There are many ways in which a teacher can become a learner again. Information received from learner ratings and feedback can be used to direct explorations of classroom strategies, techniques, and interaction as well as personal habits that may help or hinder instructional aims. Classroom action research is now widely advocated in language teaching (Allwright & Bailey, 1991; Nunan, 1988; Pennington, 1988; Widdowson & Candlin, 1990). One example of action research is the

## Learner Evaluations and Language Teaching

questioning of one's own teaching through the use of diary studies (Maurice, 1989). The purpose of such research is to enhance teacher awareness and performance by reflecting upon what is done and what goes on in the classroom.

Another example of action research for teacher development is to study specific questions of importance in the classroom. Nunan (1988) gives a few examples of such projects, including (a) seeing whether referential questions lead to more complex language than display questions; (b) getting learners to talk more and monitoring their efforts; and (c) checking whether small group activities improve the quality of learning (p. 149). Other methods of teachers being learners involve the use of systematic observation of classrooms by others, team teaching, and using video to record and analyze classroom processes (Nunan, 1988). Still others include attendance and participation in specific skill workshops, seminars, and conferences.

Whatever the method used, the importance of ongoing teacher development is obvious. ACTFL's provisional program guidelines for foreign language teacher education (ACTFL, 1988), though they deal with preservice education, seem generally applicable to most, if not all, inservice teachers as well. The guidelines focus on three general areas of development: personal, professional, and specialist. The area of personal development, which includes communication, acquisition and use of knowledge, and leadership, serves as a good example of what all of us need to work on to become the best possible teachers we can be. The sub-area of communication, which includes such matters as attuning to non-verbal cues, explaining clearly, and tailoring messages to fit different audiences, is one that is critically important, but which may take a lifetime to master.

### *4.3 Possible Problems with Teachers Being Learners*

Brock (1981) calls for caution in expecting too much of solutions generated by evaluation-based development. He notes, for example, that evaluations seem to work better with those teachers who feel that their teaching is the chief cause of student progress than it does with

teachers who feel it is all up to the students. Likewise, those who receive fair-to-fairly good ratings from learners tend to improve more than those who receive low ratings. Also, teachers who rate themselves better than students rate them are more likely to change than teachers who agree with student ratings. Brock (1981) adds that "patience and persistence are required to alter the complex set of behaviors that comprise a teacher's style" (p. 241).

#### *4.4 Cases of Teachers as Learners*

This section completes each of the three mini-cases discussed earlier.

##### *4.4.1 The ineffective teacher with little support*

Because part of the problem with this teacher's performance was due to inadequate training and support for teachers in the program, and because the teacher was sincere about improving and open to suggestion, the actions taken to build a support system helped the teacher to learn how to use materials more effectively. The personality of the teacher remained basically unchanged, but his willingness to use what others suggested and gave to him led to somewhat improved evaluations from clients. The teacher did learn.

In addition, the teachers' resource file evolved so that teachers in that branch of the company became involved in building, expanding, and improving it. Later, when the contents of the file were shared with teachers in other branches, many more teachers learned to use the materials and to support the idea of sharing. Still later, the company decided to publish a workbook of many of the worksheets; teachers who contributed were recognized for their efforts. This in turn helped to further develop a sense of pride and desire to pursue excellence among the teachers involved.

What began as one set of problems with one particular teacher was turned into an opportunity to explore ways to improve not only that teacher's work, but also, potentially, the work of every other teacher in the company.

## Learner Evaluations and Language Teaching

### *4.4.2 The course that fit the program's objectives but which generated unsatisfactory process and results.*

The teacher in this case at first seemed open to learning but later, by choosing to take a reduced role in the course, and then not being involved in the course at all, turned away from the learning opportunities in dealing with the course. The evaluator became a co-teacher in the course, by necessity, and learned much through trial-and-error about what worked and what did not work with the students. From a program and learner perspective, improvements were made in the course, but not because of the initial teacher's development.

### *4.4.3 The ineffective teacher who did not admit the problem*

In this case, the learners' continued discontent moved the evaluator to try to gently nudge the teacher toward revising how and what she did in the classroom. Her refusal to face the problem, however, led to a further isolation of the teacher in the course. The person who became her co-teacher reported that her previous well- documented tendencies persisted and that she resisted any efforts to collaborate or revise her teaching. Learner evaluations of the teacher, while extremely helpful in identifying patterns of behavior and giving insights into problem areas, could not successfully be used in this instance to persuade the teacher to improve her professional performance. From a program perspective, however, improvements were made in the course and more improvements could be made based on what was learned from the learners. Both the co-teacher's efforts and the supervisor's actions were appreciated by the learners.

## 5. Conclusion

One of the purposes of this article has been to indicate the value of using learner evaluations of teachers and courses in language programs. Learner evaluations, if designed, collected, interpreted and used properly, can have far-reaching implications on program and professional development. First, they provide valuable data to teachers about levels of student interest in the class, pace of classroom activities, rapport between teachers and learners, and other classroom concerns. This data

can help teachers to look at their teaching from another perspective, reflect on their own strengths and weaknesses, and seek ways to improve on the weak areas. Second, this data can also help supervisors to determine program strengths and weaknesses, make changes in teaching assignments according to program priorities, identify areas to be pursued in staff development activities, and provide a foundation from which to make important personnel decisions about promotion and contract renewal. Third, the professional use and follow-through actions taken as a result of such evaluations can tangibly and emphatically indicate to all concerned that the program's philosophy is client-centered and that its operations are designed and implemented with that philosophy in mind.

Another purpose of this article has been to show that the changing and combining of the roles of learner, teacher, and evaluator are key elements in whether learner evaluations are effective in improving program and instructional effectiveness. Despite widespread apprehensions among teachers, adults and university-age learners tend to be fair, consistent, and precise in their evaluations (Aleamoni, 1981). Learners *can* evaluate. In situations where learners have not evaluated their courses previously, supervisors need to prepare and explain the reasons to both the learners and the teachers to ensure that the evaluations will be seen in the proper perspective and taken seriously. For teachers to benefit from such evaluations, they must be able and willing to learn from the learners. For some, this may mean a significant shift from seeing learners as subordinates or adversaries to seeing them as clients or partners. However, if teachers need to learn and improve throughout their careers, as the professional literature suggests, then this shift is essential.

Supervisors and evaluators also need to shift roles to make the best use of learner evaluations. The teaching function of the evaluator is just as important as the judging function. Evaluators need to use triggering events to help teachers ask better questions about their teaching and to discard frozen and ineffective beliefs and behavior (McLaughlin & Pfeifer, 1988). Evaluators need to help teachers learn soon after the

## Learner Evaluations and Language Teaching

learner evaluations are in, focus on specific concerns, merit the teachers' respect, and guide in ways that teachers feel are generally supportive (McLaughlin & Pfeifer, 1988). None of these desiderata are easy, but all are important to the health of a language program.

As the mini-case studies indicate, not every effort succeeds. However, the process of having learners evaluate and of taking those evaluations seriously can help supervisors to improve a program in many ways, from changing teaching and work assignments, to choosing which areas to pursue in staff development, to making promotion and contract renewal decisions.

Many specific ways of designing and handling feedback and evaluations from learners are available to the language teaching profession. What fits one program may be not be entirely suitable for another. However, certain goals need to be met: (a) Learners should be given opportunities to give useful feedback; (b) teachers should be willing to listen to the learners; (c) evaluators and supervisors should establish an environment in which learners can contribute their ideas; (d) teachers should realize that they need to keep learning; and (e) the process should lead to more effective learning, teaching, and working relationships.

**Keith Maurice** is the assistant director of the Center for Intensive English Studies at Florida State University. He has also taught and supervised teachers in Japan and Thailand. His presentations and publications have been on business communication, communicative methodology, intercultural communication, qualitative evaluation and research, and techniques of teaching speaking.

### References

- ACTFL. (1988). ACTFL provisional program guidelines for foreign language teacher education. *Foreign Language Annals*, 21, 71-81.
- Allwright, D., & Bailey, K. (1991). *Focus on the language classroom: An introduction to classroom research for language teachers*. Cambridge: Cambridge University Press.
- Aleamoni, L. M. (1981). Student ratings in instruction. In J. Millman (Ed.), *Handbook of teacher evaluation* (pp. 110-145). Beverly Hills: Sage.
- Armiger, M. L. (1981). The political realities of teacher evaluation. In J. Millman

- (Ed.), *Handbook of teacher evaluation* (pp. 292-302). Beverly Hills: Sage.
- Breen, M. P., & Candlin, C. N. (1980). The essentials of a communicative curriculum in language teaching. *Applied Linguistics*, 12, 89-112.
- Bridges, E. M. (1986). *The incompetent teacher: The challenge and the response*. Philadelphia: The Falmer Press.
- Brock, S. C. (1981). Evaluation-based teacher development. In J. Millman (Ed.), *Handbook of teacher evaluation* (pp. 229-243). Beverly Hills: Sage
- Cronbach, L. J. (1983). *Designing evaluations of educational and social programs*. San Francisco: Jossey-Bass.
- Christison, M. A., & Krahnke, K. J. (1986). Student perceptions of academic language study. *TESOL Quarterly*, 20, 61-82.
- Darling-Hammond, L., Wise, A. E., & Pease, S.R. (1983). Teacher evaluation in the organizational context: A review of the literature. *Review of Educational Research*, 53, 285-327.
- Doyle, K. O., Jr. (1983). *Evaluating teaching*. Lexington, MA: D. C. Heath.
- Fanselow, J. F. (1987). *Breaking rules: Generating and exploring alternatives in language teaching*. New York: Longman.
- Guba, E. G., & Lincoln, Y. S. (1983). *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches*. San Francisco: Jossey-Bass.
- Hutchinson, T., & Waters, A. (1987). *English for specific purposes: A learning-centred approach*. Cambridge: Cambridge University Press.
- Lewkowicz, J. & Moon, J. (1985). Evaluation: A way of involving the learner. In C. Alderson (Ed.), *Evaluation: Lancaster practical papers in English language education* (pp. 45-80). Oxford: Pergamon.
- Loew, H. Z. (1979). Modifying the program and providing for change. In J.K. Phillips (Ed.), *Building on experience—Building for success* (pp. 271-304). Skokie, IL: National Textbook Co.
- Maurice, A. (1989). Diary studies as a teacher awareness tool. *Guidelines*, 11 (2), 88-98.
- McLaughlin, M. W., & Pfeifer, R. S. (1988). *Teacher evaluation: Improvement, accountability, and effective learning*. New York: Teachers College Press.
- McNeil, J. D. (1981). Politics of teacher evaluation. In J. Millman (Ed.), *Handbook of teacher evaluation* (pp. 272-291). Beverly Hills: Sage.
- Murphy, D. F. (1985). Evaluation and language teaching: Assessment, accountability, and awareness. In C. Alderson (Ed.), *Evaluation: Lancaster practical papers in English language education* (pp. 1-18). Oxford: Pergamon.
- Nunan, D. (1988). *The learner-centred curriculum*. Cambridge: Cambridge University Press.
- Omaggio, A. C., Eddy, P. A., McKim, L. W., & Pfannkuche, A. V. (1979). Looking at the results. In J. K. Phillips (Ed.), *Building on experience—Building for success* (pp. 233-270). Skokie, IL: National Textbook Co.

## Learner Evaluations and Language Teaching

- Pennington, M. C. (1989). Faculty development for language programs. In R.K. Johnson (Ed.), *The second language curriculum* (pp. 91-110). Cambridge: Cambridge University Press.
- Pennington, M. C., & Young, A. L. (1989). Approaches to faculty evaluation for ESL. *TESOL Quarterly*, 23, 619-646.
- Richards, J. C. (1990). The teacher as self observer: Self-monitoring in teacher development. In *The language teaching matrix* (pp. 118-143). Cambridge: Cambridge University Press.
- Scriven, M. (1981). Summative teacher evaluation. In J. Millman (Ed.), *Handbook of teacher evaluation* (pp. 244-271). Beverly Hills: Sage.
- Sharp, A. (1990). Staff/student participation in course evaluation: A procedure for improving course design. *ELT Journal*, 44, 132-137.
- Sheal, P. (1989). Classroom observation: Training the observers. *ELT Journal*, 43, 92-103.
- Smith, B. O. (1977). Teacher education as perennial. In J. F. Fanselow & R. L. Light (Eds.), *Bilingual, ESOL, and foreign language teacher preparation: Models, practices, issues* (pp. 179-190). Washington, D.C.: TESOL.
- Stenhouse, L. (1975). *An introduction to curriculum research and development*. London: Heinemann.
- Stiggins, R. J., & Duke, D. (1988). *The case for commitment to teacher growth: Research on teacher evaluation*. Albany: State University of New York Press.
- Waters, A. (1987). Participatory course evaluation in ESP. *English for Specific Purposes*, 6, 59-68.
- Widdowson, H. G., & Candlin, C. N. (1990, March). *Teacher education through action research*. Paper given at TESOL '90 Convention, San Francisco.
- Williams, M. (1989). A developmental view of classroom observations. *ELT Journal*, 43, 85-91.

**The  
Canadian  
Modern  
Language  
Review**

**La Revue  
canadienne  
des langues  
vivantes**

---

**Anthony S. Mollica**  
**Editor**

Linguistic and pedagogical articles, book reviews, current advertisements and other material of interest to teachers of French, English, German, Italian and Spanish, as second languages at all levels of instruction as well as to teachers of Heritage languages.

**Subscription rates:**

- Individual \$25.00    Sustaining \$ 50.00   
 Institutions \$35.00    Patrons        \$100.00

(U.S.A. and Overseas  
Subscriptions in U.S. Funds)

---

**Canada's Voice in Language Teaching and Learning**

Founded in 1944.

Published regularly in October, January, March and May. Occasionally, a special supplementary issue is also published.

Cheques or money orders payable to *The Canadian Modern Language Review/La Revue canadienne des langues vivantes* should be sent to - Managing Editor, CMLR/RCLV, 237 Hellems Avenue, Welland, Ontario L3B 3B8 (Canada).

Sample copy of the journal will be sent on request.



-----

Enclosed please find my cheque or money order for \$ \_\_\_\_\_ for a one-year subscription to *The Canadian Modern Language Review/La Revue canadienne des langues vivantes (CMLR/RCLV)*, 237 Hellems Avenue, Welland, Ontario L3B 3B8 (Canada).

Please begin my subscription with the October , January , March , May  issue.

please print

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

Address: \_\_\_\_\_

\_\_\_\_\_

City: \_\_\_\_\_ Prov./State: \_\_\_\_\_

Postal/Zip Code: \_\_\_\_\_

Signature: \_\_\_\_\_

Date: \_\_\_\_\_

Please return this portion with your payment. Thank you.

## **EFL Learners Talking to Each Other: Task Types and the Quality of Output**

**Atsuko Ushimaru**  
*Obirin University*

The notion of “comprehensible output,” or language production pushed toward the target norm (“pushed output”), is relevant in the EFL context, where learners of English typically interact with other learners. The study reported in this article investigated whether interlocutors in nonnative speaker-nonnative speaker (NNS-NNS) interactions reformulate their utterances in more grammatical language in response to signals of incomprehension, as they do in talking to native speakers. The study observed how NNSs behave linguistically under different task conditions, with a focus on their grammaticality, incomprehension signals, and subsequent reformulations. It was found that pushed output does occur to some extent in NNS-NNS interaction, but this did not coincide with the degree of overall grammaticality.

### **EFL 学習者同士の会話におけるアウトプットの質の分析**

日本のような EFL 環境では、同じ母国語を持つ学習者同士で英語を話すことが多い。その際にも、ネイティブ・スピーカーと話すときと同じように、相手が理解できなかった発話をより文法的には正しく言い直すという行動 (pushed output) が見られるだろうか。本稿は、3つの異なる作業における学習者同士の会話で、英語がどのような形をとるか、文法的正しさ、理解できなかったときの合図、それに対する反応、といった点から分析した研究の報告である。この研究では、EFL 学習者同士の会話でも pushed output はある程度起こるが、その量と発話全体の文法的正確さの数は一致しないという結果が得られた。

### **1. Introduction**

Where do language learners stand as interactional partners for other learners? In the current theories of second language acquisition, inter-

actional modifications, such as confirmation and comprehension checks and clarification requests, are regarded as requisite to the development of learners' receptive and productive capacities. These have been found to exist not only in interaction between native speakers (NSs) and learners, but also between learners—and to serve similar functions (e.g., Chaudron, 1983; Duff, 1986; Gaies, 1983; Gass & Varonis, 1985; Long, 1985; Porter, 1986; Seliger, 1983; Varonis & Gass, 1982), thus justifying nonnative (NNS) speakers as legitimate input providers.

In recent years, focus has been directed also to the role of learners' output. As more and more data are gathered through immersion studies (e.g., Lambert & Tucker, 1972; Plann, 1977; Swain, 1985) and from studies of fossilized learners (e.g., Schmidt, 1983), it is becoming clear that input is not the sufficient condition for reaching native-speaker level. Swain (1985) found, through research on French immersion students' linguistic competence, that although comprehensible input had been abundantly available to NNSs for as long as seven years, their grammatical performance was not equivalent to that of NSs. Swain concludes that the impact of input and interaction on grammatical development has been overstated in previous research, and that comprehensible input alone is not enough for grammatical competence to develop adequately. Whereas Krashen (1981) saw the role of output as only that of generating comprehensible input, Swain claims that it provides the opportunity for learners to meaningfully use their linguistic resources, and that:

In order for native-speaker competence to be achieved, . . . the meaning of "negotiating meaning" needs to be extended beyond the usual sense of simply "getting one's message across." Simply getting one's message across can and does occur with grammatically deviant forms and sociolinguistically inappropriate language. Negotiating meaning needs to incorporate the notion of being pushed toward the delivery of a message that is not only conveyed, but that is conveyed precisely, coherently, and appropriately (pp. 248-249).

## Quality of Learner Output in EFL

Swain calls this sort of output “comprehensible output,” meaning that the message is made comprehensible and acceptable to the NSs, and is therefore target-like.

Subsequent research by Pica and her associates (Pica, 1987; Pica, Holliday, Lewis, & Morgenthaler, 1989), which operationalized the construct of “comprehensible output” to provide it with direct empirical grounding, tested whether interaction with NSs actually made the learners’ output more target-like. Pica (1987) found that in an open-ended discussion task this did happen, although the learners were given only limited opportunities to reformulate their utterances which were incomprehensible to the NSs. The results from Pica et al. (1989), in which three different tasks were used, suggest that the type of task could, to a certain extent, influence the number of opportunities that NSs provide learners to reformulate their imperfect utterances. Further, the type of interactional modifications (specifically clarification requests and confirmation checks, which the researchers called “incomprehension signals”) had a relationship with whether or not learners actually responded to the opportunities by making their subsequent utterances more target-like. The question now is whether learners talking to other learners can also be moved somehow to make their output more grammatical through reformulation of utterances.

### 2. The Present Study

The notion of “pushed output”—language production pushed toward the target norm—is especially relevant in an EFL context such as Japan. In an environment where exposure to the target language outside the classroom (and very often inside the classroom as well) is limited, the learners often have no choice but to practice the second language (L2) with other learners who share the same first language (L1). Previous research suggests that the frequency of interactional modification sequences is lower among interlocutors of the same L1 background (Doughty & Pica, 1986). It is also feared that the amount of grammatical input might be restricted by a steady diet of group activities (cf., Pica & Doughty, 1985), and that the development of

“classroom dialects”—non-target varieties of the L2 intelligible only among learners in the same classroom—is possible, such as is observed by Hatch (1978), Lambert and Tucker (1972), Plann (1977), Selinker (1975), and Taylor (1982).

### *2.1 Purpose of the Study: Research Questions*

The general purpose of this paper is to identify some of the conditions under which learners talking to other learners in the EFL context are “pushed,” if at all, to produce, when given signals by their partners, output that is more grammatical and target-like. The term “target-like” rather than “comprehensible” will be used in this paper, for it is easy for interlocutors sharing the same L1 to understand each other’s interlanguage (Porter, 1986), and comprehensibility of output, unlike in NS-NNS interaction, does not guarantee its approximation to the target norm.

The variables that have been focused on in previous research as influencing the quality of NNS output, besides interlocutor ethnicity (e.g., Duff, 1986; Porter, 1986), include: task type (e.g., Duff, 1986; Gass & Varonis, 1985; Pica et al., 1989), gender (e.g., Gass & Varonis, 1985; Pica et al., 1989), and proficiency (e.g., Pica & Doughty, 1985; Porter, 1986). From among these, task type is taken up in this study as the one that is most relevant to the school setting, where the teaching of English typically takes place. It can fairly realistically be, and actually is, manipulated, while the human factors mentioned above are often beyond control of the classroom teachers. Adapting the design in Pica et al. (1989), this study used three types of tasks, namely information-gap (one participant has all the necessary information); jig-saw (each participant has some information that the others need); and open-ended discussion (the participants talk freely on a given topic)—to create situations with varying degrees in the precision of information required. The highest degree of precision would be called for in the information-gap task, less in the jig-saw task, and the least in the open-ended discussion.

The present study (a) measured the overall grammaticality of the language, which was Swain’s (1985) original concern, and (b) observed

## Quality of Learner Output in EFL

the learners' behavior in talking to other learners sharing the same L1 under different task conditions, in terms of incomprehension signals and subsequent reformulations. The following are the specific research questions asked:

1. Will the overall grammaticality of the language produced differ across tasks?
2. Will the number of incomprehension signals, that is, clarification requests and confirmation checks, differ across tasks?
3. Will clarification requests and confirmation checks be made in different proportions across tasks?
4. Will the amount of reformulations made in response to incomprehension signals differ across tasks?
5. Which incomprehension signal will be responded to more with reformulations?

### *2.2 Design*

#### *2.2.1 Subjects*

The subjects for this study were 12 high school seniors (ages 17-18) at a private girls' school in Japan, enrolled in an elective class of English conversation which met for three 45-minute sessions a week and mainly engaged the students in communicative activities. They were all Japanese nationals sharing the same language background, with proficiency levels ranging from lower to higher intermediate. At school they had had English for six years, four hours a week, in teacher-fronted classes with grammatical syllabuses. They had all taken part in short-term (one month) homestay programs in English-speaking countries during the previous two years, but had not lived abroad for extended periods. Besides this elective class, the subjects were taking several reading and grammar classes of English, all of which were in the teacher-fronted, lockstep format. For this study, the subjects were divided randomly into six pairs.

#### *2.2.2 Tasks*

The following list describes each of the actual tasks:

1. **Information-gap.** One of the interlocutors described a drawing of a tennis racket and directed the other to draw the same, providing detailed information as to its size, shape, and location on the paper. All the pairs had the same drawing, so that difference in the pictures would not influence the language produced.
2. **Jig-saw.** The task used a 10-frame picture sequence about a man who meets with repeated mishaps as he walks home. One of the interlocutors had five of the frames in proper sequence, and the rest were randomly arranged in a separate place together with seven extra pictures that had some features in common with one or more of the correct pictures. The other interlocutor had the other five correct frames in place, with five correct and seven extra pictures randomly arranged. The seven extra pictures were the same for both interlocutors. They were to try to reconstruct the original sequence by exchanging information.
3. **Open-ended discussion.** The subjects were to discuss their favorite pastime, a topic chosen for its personal nature—one which they would find easy to relate to. They were allowed to deviate from the topic during the course of the discussion.

Through pretesting with two pairs of students comparable to the subjects in the present study, it was found that at their level the tasks were challenging but not too difficult.

### *2.2.3 Data collection*

The tasks were done in the researcher's office. The pairs were called in one by one, and each pair worked in isolation from the others. The subjects sat facing each other, and during the information-gap and jig-saw tasks a cardboard screen was placed between them so that they could not see their partners' pictures. They worked on all three tasks at one sitting but in different sequences, in order to avoid contamination of data by practice effect and/or fatigue. All the interactions, which were recorded on audio tape, were cut off after 15 minutes, whether or not the tasks had been completed.

## Quality of Learner Output in EFL

### 2.3 Analysis

The first five minutes were ignored in the analysis, giving time for the speakers to “get into gear,” so to speak. Each interaction thus resulted in a 10-minute sample. All the samples were transcribed and coded for c-units and their grammaticality, incomprehension signals, and reformulations in response to these signals.

A **c-unit** was defined as a “word, phrase, or sentence that in some way contributes pragmatic or semantic meaning to a conversation” (Duff, 1986, p. 153), and

a segment of NNS speech was not disqualified as a c-unit because it lacked or included incorrectly the copula, the impersonal pronoun *it*, an auxiliary verb, prepositions, articles, or inflectional morphology. (Brock, 1986, p. 52)

The measure of grammaticality adopted was that of Pica and Doughty (1985): the proportion of grammatical (i.e., containing no errors in morphology, syntax, or lexis) c-units out of the total number of c-units in each sample. Single-word c-units were excluded from this calculation. Those consisting solely of Japanese words were eliminated from the calculation of grammaticality, but were included in the analysis of incomprehension signals.

Incomprehension signals were identified on the following basis (cf., Long, 1980; Pica et al., 1987):

1. Clarification requests: Moves intended to elicit clarification of the preceding utterance, through *wh-*, *yes-no*, *polar*, *disjunctive*, *uninverted*, and *tag* questions and statements such as “I don’t understand” and “Please repeat.”
2. Confirmation checks: Moves immediately following the previous speaker’s utterance to seek confirmation that it has been understood correctly, through repetition of all or part of the utterance. Reformulations were those responses that indicated any improvement at all toward the target norm compared to the original

utterances that had triggered the incomprehension signals, and were therefore not necessarily completely grammatical.

### 3. Results and Discussion

The results of data analysis indicate that L2 learners' output varies to a certain extent in quality depending on the type of task used to elicit the language.

Research Question one asked whether there would be a difference in the overall grammaticality of output across tasks. As shown in Table 1, it was found that the language produced during the open-ended discussion was the most grammatical of the three, and that made during the information-gap task was the least grammatical. The order is reversed in the case of the number of incomprehension signals (Research Question two), the subjects giving the largest number of signals when engaged in the information-gap task, fewer in the jig-saw task, and fewest in the open-ended discussion.

**Table 1**  
**Comparisons of grammatical output and incomprehension signals in c-units across task**

Task	Total	Grammatical output		Incomprehension signals	
	n	n	weighted*	n	weighted *
IGT**	712	189	251.12	121	160.93
JST**	946	312		101	
DSC**	780	337	408.72	72	78.12
IGT vs. JST vs. DSC		$X^2 = 38.99$		$X^2 = 26.45$	df = 2. $p < .05$
IGT vs. JST***		$X^2 = 16.55$		$X^2 = 18.60$	df = 1. $p < .05$
IGT vs. DSC***		$X^2 = 38.07$		$X^2 = 23.79$	df = 1. $p < .05$
JST vs. DSC***		$X^2 = 22.32$		$X^2 = 8.99$	df = 1. $p < .05$

\*IGT: information-gap task; JST: jig-saw task; DSC: open-ended discussion

\*\*Frequencies are corrected for the unequal amount of total output.

\*\*\*Yates correction factor for continuity is applied.

## Quality of Learner Output in EFL

These results seem to reflect the nature of the three tasks in terms of the requirement for conveyance of information and the precision of the information conveyed (cf., Pica et al., 1989). In the open-ended discussion, there was no conclusion that had to be reached and no final configuration to be constructed; therefore the subjects were able to choose to say only what they wanted or knew how to say. (In fact, all the pairs deviated from the originally assigned topic after a time.) In the case of the information-gap task, on the other hand, in order to direct the partner to draw the picture precisely, the speaker had to try to give the information whether or not she knew how. The jig-saw task stands in the middle: Although conveyance of information was required to reach the goal, the speakers did not have to explain everything, since, in this task, the missing information was divided between the interlocutors and they had all the pictures being described by their partners. This is likewise reflected in the number of incomprehension signals made in the execution of each task. Pressure to clarify ambiguities is greatest for an information-gap task, whereas in the case of an open-ended discussion, incomprehension of certain information does not necessarily get in the way of overall interaction. Pica (1987) had noted that when NS-NNS pairs engaged in open-ended discussion, only a limited number of opportunities were given to the NNSs to reformulate utterances which were incomprehensible to the NSs; a parallel result was observed in NNS-NNS pairs in the present study.

No significant difference was found in the proportion of clarification requests and confirmation checks, nor in the total number of reformulated responses across the three tasks (Research Questions three and four). But of the two types of incomprehension signals, clarification requests led to more reformulations than confirmation checks did overall, although the information-gap task alone reflected this result (Research Question five). Tables 2 and 3 give more details of these results.

Simply stated, incomprehension signals were given most frequently on the information-gap task, and clarification requests led to more reformulated responses than confirmation checks, which in turn oc-

cured most on the information-gap task. It would, however, be rash to associate this directly with the task's effect on the learners' development in the L2. Some reservations should be kept in mind: First, the occurrence of a larger number of incomprehension signals (opportunities for reformulation) obviously does not in itself guarantee higher overall grammaticality in the language produced. The information-gap

**Table 2**  
**Comparisons of type of incomprehension signal and reformulated responses in c-units across task**

Task	Incomprehension signals					Reformulated responses	
	Total	CL*		CF*		n	w**
	n	n	w**	n	w**		
IGT	121	81		40		29	
JST	101	51	61.2	50	60	26	31.2
DSC	72	41	68.88	31	52.8	14	23.52
<b>Total</b>	<b>173</b>			<b>121</b>			
IGT vs. JST vs. DSC			$X^2 = 2.83$			$X^2 = 1.12$	df = 2, ns
IGT vs. JST***			$X^2 = 2.53$			$X^2 = 0.29$	df = 1, ns
IGT vs. DSC***			$X^2 = 1.47$			$X^2 = 0.55$	df = 1, ns
JST vs. DSC***			$X^2 = 1.08$			$X^2 = 0.82$	df = 1, ns

\*CL: clarification requests; CF: confirmation checks  
 \*\*w: weighted (Frequencies are corrected for the unequal total number of incomprehension signals.)  
 \*\*\*Yates correction factor for continuity is applied.

task gave rise to more incomprehension signals, especially clarification requests—the sort of interaction that “pushes” the production toward the target norm—than the other two tasks did, but the interaction that occurred during this task also had the lowest grammaticality. On the other hand, although the pairs did not engage in much negotiation of meaning in the open-ended discussion, their interaction maintained the best-formed production. One is inclined to debate between the merits of opportunities for output to be “pushed,” and the demerits of NNSs feeding each other malformed language (cf., Pica & Doughty, 1985).

## Quality of Learner Output in EFL

**Table 3**  
**Comparison of reformulations in response to each**  
**type of incomprehension signal in c-units**

Task	CL n	RCL* n	CF n	RCF* n	RCF* weighted**
IGT	81	24	40	5	10.13
JST	51	15	50	11	11.22
DSC	41	11	31	3	3.97
<b>Total</b>	<b>173</b>	<b>50</b>	<b>121</b>	<b>19</b>	<b>27.17</b>
<b>Total***</b>	$X^2 = 6.16$ df = 1, $p < .05$				
<b>IGT***</b>	$X^2 = 5.64$ df = 1, $p < .05$				
<b>JST***</b>	$X^2 = 0.29$ df = 1, ns				
<b>DSC***</b>	$X^2 = 2.43$ df = 1, ns				

\*RCL: reformulated responses to clarification requests; RCF: reformulated responses to confirmation checks  
 \*\*Frequencies are corrected for the unequal number of incomprehension signals.  
 \*\*\*Yates correction factor for continuity is applied.

At the same time, although it was found that the subjects, when given incomprehension signals, did reformulate their utterances toward the target norm, only a very few of these signals actually led to such reformulated responses. Other types of responses included: incorrect or irrelevant responses, repetition of incomprehension signals that contained grammatical errors, mere acknowledgment of signals (i.e., yes, uh-huh), and repetition of triggers which were grammatical to begin with (the partner did not understand the utterance although it was correct). Some of the signals received no responses whatsoever. It seems that a considerably large portion of incomprehension signals—initiatives to negotiation of meaning—are wasted: The functions, if any, of those signals that are given and received but fail to be responded to overtly are still unclear (cf., Pica et al., 1989).

#### 4. Conclusion

The present study, motivated by the notion of “comprehensible output” and the need for “pushed” output, investigated the quality of language produced between NNS interlocutors under different task conditions. Results indicate that “pushed” output does occur to some extent in NNS-NNS interaction, as it does in NS-NNS interaction, that is, incomprehension is dealt with by making subsequent utterances more grammatical.

This study is significant in that it was conducted in an EFL environment where exposure to the target language is limited and the learners’ L2 interlocutors are most likely peers who share the same L1, so that this notion is especially relevant. It is significant also because it showed that in discussing the occurrence of incomprehension signals and reformulated responses one should probably take into consideration the degree of overall grammaticality as well. It is limited, however, because it used only one aspect of task types (i. e., precision of information required) as the variable influencing output—the results are examples of what happens in one specific situation. Other dimensions of task types should probably be explored, as well as conditions other than task types that bring about differences in the quality of output.

More studies of this sort need to be conducted so long as the issues referred to in the previous section remain unsolved: For instance, it would not be wise at this point to dismiss the role of confirmation checks (Gass & Varonis, 1985; Pica et al., 1989); and it is possible that one key to second language acquisition is to become aware that one’s utterance has been incomprehensible and to make efforts to reformulate it (Ellis, personal communication), no matter what the overall shape of the resulting language. Conclusions concerning the effectiveness of NNS interactional partners as compared to NSs in “pushing” output must definitely wait.

This is a cross-sectional study which momentarily captures the effects of whatever language-learning experiences learners have had in

## Quality of Learner Output in EFL

the past, and not necessarily what would happen in the future if the type of activities in the research project were to be continued. The incorporation of language input is never immediate (Lightbown, 1983), and therefore longitudinal studies are also critically needed.

*This study was funded by the 1990 JALT Grant for Research in Language Teaching and Learning. An earlier version of this paper was presented at the 16th Annual JALT International Conference in Omiya, November 23-25, 1990.*

**Atsuko Ushimaru** is an assistant professor at Obirin University and a doctoral candidate in TESOL/Applied Linguistics at Temple University Japan. Her major research interest is classroom second language acquisition in the EFL context.

### References

- Brock, C. A. (1986). The effects of referential questions on ESL classroom discourse. *TESOL Quarterly*, 20, 47-59.
- Chaudron, C. (1983). Foreigner talk in the classroom—An aid to learning? In H. W. Seliger & M. H. Long (Eds.), *Classroom oriented research in second language acquisition* (pp. 127-145). Rowley, MA: Newbury House.
- Doughty, C., & Pica, T. (1986). "Information gap" tasks: Do they facilitate second language acquisition? *TESOL Quarterly*, 20, 305-325.
- Duff, P. A. (1986). Another look at interlanguage talk: Taking task to task. In R. R. Day (Ed.), *Talking to learn* (pp. 147-180). Rowley, MA: Newbury House.
- Gaies, S. (1983). The investigation of language classroom processes. *TESOL Quarterly*, 17, 205-217.
- Gass, S. M., & Varonis, E. M. (1985). Task variation and nonnative/nonnative negotiation of meaning. In S. M. Cass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 149-161). Rowley, MA: Newbury House.
- Hatch, E. (1978). Apply with caution. *Studies in Second Language Acquisition*, 2, 123-143.
- Hatch, E. (1983). *Psycholinguistics: A second language perspective*. Rowley, MA: Newbury House.
- Krashen, S. D. (1981). The "fundamental pedagogical principle" in second language teaching. *Studia Linguistica*, 35, 50-70.
- Lambert, W. E., & Tucker, R. (1972). *Bilingual education of children: The St. Lambert experiment*. Rowley, MA: Newbury House.

- Lightbown, P. M. (1983). Exploring relationships between developmental and instructional sequences in L2 acquisition. In H. W. Seliger & M. H. Long (Eds.), *Classroom oriented research in second language acquisition* (pp. 217-245). Rowley, MA: Newbury House.
- Long, M. H. (1980). *Input, interaction, and second language acquisition*. Unpublished doctoral dissertation, University of California at Los Angeles.
- Long, M. H. (1985). 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.
- Pica, T. (1987). Interlanguage adjustments as an outcome of NS-NNS negotiated interaction. *Language Learning*, 38, 45-73.
- Pica, T., & Doughty, C. (1985). Input and interaction in the communicative language classroom: a comparison of teacher-fronted and group activities. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 115-132). Rowley, MA: Newbury House.
- Pica, T., Holliday, L., Lewis, N., & Morgenthaler, L. (1989). Comprehensible output as an outcome of linguistic demands on the learner. *Studies in Second Language Acquisition*, 11, 63-90.
- Pica, T., Young, R., & Doughty, C. (1987). The impact of interaction on comprehension. *TESOL Quarterly*, 21, 737-758.
- Plann, S. (1977). Acquiring a second language in an immersion classroom. In H. D. Brown (Ed.), *On TESOL '77* (pp. 213-225). Washington, D.C.: TESOL.
- Porter, P. A. (1986). How learners talk to each other: Input and interaction in task-centered discussions. In R. R. Day (Ed.), *Talking to learn* (pp. 200-222). Rowley, MA: Newbury House.
- Schmidt, R. (1983). Interaction, acculturation, and the acquisition of communicative competence. In N. Wolfson & E. Judd (Eds.), *Sociolinguistics and language acquisition* (pp. 137-174). Rowley, MA: Newbury House.
- Seliger, H. W. (1983). Learner interaction in the classroom and its effect on language acquisition. In H. W. Seliger & M. H. Long (Eds.), *Classroom oriented research in second language acquisition* (pp. 246-267). Rowley, MA: Newbury House.
- Selinker, L. (1975). Interlanguage. *IRAL*, 10, 209-231.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.
- Taylor, B. P. (1982). In search of real reality. *TESOL Quarterly*, 16, 29-42.
- Varonis, E., & Gass, S. (1982). The comprehensibility of non-native speech. *Studies in Second Language Acquisition*, 4, 41-52.

# NEW!

## The Heinemann **TOEFL**<sup>®</sup> Preparation Course

*"The most complete TOEFL book on the market. . ."*

Provides successful test-taking  
TOEFL strategies



Improves your language skills *and*  
test scores



Covers the Test of Written English

**M. Kathleen Mahnke & Carolyn B. Duffy**



Consultant: Jacqueline Flamm

"The most complete TOEFL book on the market. A real find for the classroom teacher as well as for the self-study student. A written guarantee of success."

*Martin Howard, Director,  
Intensive English Program,  
Indiana University*

For further information on all HEINEMANN publications, please contact:



**HEINEMANN**  
**INTERNATIONAL**

Vaughan Jones, Japan Manager, Shin Nichibo Building, 1-2-1 Sarugakucho,  
Chiyoda-Ku, Tokyo 101 Telephone: 03 3294-0791 Fax: 03 3294-0792



SOUTHEAST ASIAN  
MINISTRY OF EDUCATION ORGANIZATION  
Regional Language Centre

## MA in Applied Linguistics

A New MA to be completed through course work and dissertation commencing July 1992.

### Entry Requirements

Applicants for this 15 month full-time course must have at least 2 years of recognised teaching experience and a first degree from a recognised university.

### Structure and Contents

Course work comprises the following core courses:

1. Linguistics in Language Teaching
  2. Language Acquisition
  3. Language Teaching and Testing
  4. Curriculum Design and Implementation
  5. Language Research Methodology
- and four elective courses.

### Course Requirements

MA candidates have to sit for written examinations in the core courses (except for Language Research Methodology) and submit a 20,000 word dissertation.

For brochure and application form, please phone 7379044 ext 310 or fax the coupon below:

The Registrar  
SEAMEO Regional Language Centre  
30 Orange Grove Road  
Singapore 1025  
Fax: 7342753

Closing date: 04 April 1992

### MA in Applied Linguistics

Name : \_\_\_\_\_ Age: \_\_\_\_\_

Address: \_\_\_\_\_  
\_\_\_\_\_

Position: \_\_\_\_\_

Tel: (O) \_\_\_\_\_ (R) \_\_\_\_\_

## Point-to-Point

### **A Response to “Communicative Competence and the Japanese Learner”**

In *JALT Journal* 13/2, Rod Ellis published an article, “Communicative Competence and the Japanese Learner,” in which he has advocated modest proposals for the improvement of English language teaching in schools in Japan. In doing so, he shows some awareness of the constraints on what is feasible; nevertheless, his rationale, both explicit and implicit, does not withstand scrutiny.

His abstract begins with the statement that high schools are now “embracing communicative approaches.” This, unfortunately, misrepresents present conditions. Despite Mombusho’s efforts to encourage a more communicative methodology, it is quite inappropriate to use the “embracing” metaphor to describe the situation in the schools which is still characterised by an extreme form of grammar translation methodology and a lack of oral activities. One must, therefore, regard as suspect Ellis’s initial premise and, consequently, doubt the direct relevance of much of the discussion on the range of competences, for it is only grammatical competence which is the aim of schools. This is not, however, to deny the academic interest of that particular part of the article.

It is surely axiomatic that before proposing modification to an existing system, one must take into account the essential characteristics thereof. To paint a picture of English language teaching rife (“with increasing passion” [p. 104] in Ellis’s words) with communicative methodology creates a dangerously false impression. Nearer the truth (and I derive this information from, among other sources, AET’s, Japanese teachers, and observation of classes in schools) is a situation characterised by:

1. A programme dominated by a stringent examination system demanding excessive formal grammatical knowledge, often of the arcane variety, expressed largely in written form. (Ellis initially demonstrates his awareness of this but then fails to take it into considera-

tion.)

2. Teachers, students, and parents who seemingly give priority to the principles of this system and regard the conversational activities arising from communicative methodology as not serious work. (This is partly based on a survey of attitudes of over 1,000 students in the Tottori area as part of an inquiry into perceptions related to the AET programme being carried out by two colleagues and myself.)

Given this situation, it makes little sense to offer a solution based on the assumption that schools are now in the fond embrace of communicative methods. Nevertheless, for the sake of argument, let us assume that there is a degree of validity in Ellis's premise and examine his proposals in the light thereof. He proposes that:

1. Minimally, the focus should be placed on teaching formulae and a kit of rules for adapting them to contextual requirements.
2. These may best be taught "through problem-solving activities designed to raise learners' consciousness<sup>1</sup> about linguistic and sociolinguistic features of English."

As to [1], Ellis is not particularly forthcoming as to the substance of his proposal. It is, therefore, difficult to evaluate it. However, as expressed, it does not appear to be radically different from many of the current text books based on an inductive approach. Therein, there is a marked tendency to concentrate on forms which might be regarded as formulaic, particularly in the case of functions, and then later to introduce rules to explicate the patterns of the formulae.<sup>2</sup>

Though neither original nor particularly revolutionary, Ellis's proposal does have the virtue of not setting impossible goals. It may even in the present system have its place in the first and second years of junior high school where there is apparently greater possible flexibility in methodology. However, despite this, implementation would still cause problems as it would not ensure coverage of the syllabus items necessary for future levels when the all-pervasive influence of examination preparation dominates all classroom activity. This raises serious doubts concerning the feasibility of what Ellis advocates. Before being worthy

## Point-to-Point

of serious consideration, there would have to be a demonstration of a means of solving these problems.

In his second proposal, Ellis suggests that his aims in [2] above “can best be achieved through problem-solving activities designed to raise learners’ consciousness about linguistic and sociolinguistic features of English.” An implied claim entailed herein is that a problem-solving activity which aims to discover some linguistic or sociolinguistic rule is more effective than any other available method in terms of understanding and retention. This would be a strong claim to make.

It is, however, based on an intuitively seductive argument for I suppose we all may feel instinctively that discovering something for oneself is preferable to being told it. Nevertheless what is intuitive about the learning process is neither necessarily true nor appropriate for formal classroom instruction.<sup>3</sup> Relying on one’s intuition to decide personal actions is acceptable. However, when it is a question of proposing a wholesale change for a vast school system such as Japan’s as is entailed in Ellis’s proposal, one would hope that there is some form of empirical support such as, for example, a successful pilot scheme. Unfortunately, no such support is offered. This is understandable as it does not exist.<sup>4</sup> However, this IS an empirical question as Rutherford and Sharwood Smith point out in their introduction to a collection of articles on the teaching of grammar some of which are related to consciousness raising. They take the position that at this time the idea of CR should be posed for the purpose of “the stimulation of rational inquiry and not for the purpose of pushing premature decisions about how to teach languages” (1988, p. 7). Ellis, not heeding such advice, proposes the implementation of a methodology bereft of empirical support.

In spite of this, once again let us ignore a fundamental objection to Ellis’s proposals and examine the problems of implementation. Given space limitations, I will discuss just two objections. (Cf. Sheen, 1990, for a lengthier discussion of this issue.) In the first place, one of the professed advantages of the approach is the using of the target language in the problem-solving. Now, I may be over-pessimistic, but I simply

cannot conceive of students in most English classes in Japanese schools being capable of conducting a discussion of grammatical and sociolinguistic problems of the type suggested by Ellis in English. They would surely have to do so in Japanese, thus negating one of the avowed advantages of the approach.

My second objection is related to the ability required in problem solving. This is very much intelligence-related. Consequently, in a class of forty mixed ability students, only a small subset will be able to solve the problem. Most will perform a passive role and will ultimately have to be instructed in the rule, thus negating another advantage of the approach, that of the active involvement in a discovery procedure. Of course, one might argue that even this non-participatory role might be of some benefit. However, this would have to be demonstrated with empirical support for it appears to be an extremely tenuous argument.

I have objected in this response to an applied linguist of deserved reputation and influence proposing overall modifications to the school teaching of English which are based on an initial faulty premise and for the effectiveness of which there is no empirical support. In conclusion, I would like to broaden this to appeal to applied linguists as a whole. I would suggest that as applied linguists, we should limit the discussion of our ideas for major changes in methodology to an audience of fellow applied linguists until such time as there is empirical support and ample practical justification for the implementation of those ideas. This, to repeat Rutherford and Sharwood Smith (*ibid*), will go some way to preventing the making of "premature decision about how to teach languages."

#### Notes

<sup>1</sup>It was Sharwood-Smith(1981) who first introduced this idea, terming it "consciousness raising." However, White, Spada, Lightbown, and Ranta (1991, p. 417) report that in a 1991 article in *Second Language Research*, he proposes that in future it should more appropriately be termed "input enhancement."

<sup>2</sup>Ellis gives in his appendix a sample text taken from the work of other authors which presumably contains examples of his proposed formulae. If this is the case, it is somewhat puzzling for two reasons:

## Point-to-Point

- a) It is odd that he gives examples not created by himself. I would have thought that given the overall changes he was proposing he would have had examples which he had already piloted on a sample population.
- b) The examples are really quite complex. They, therefore, cannot possibly be an exponent of the simple formulae and a “kit of rules” he proposes. It is, therefore, somewhat difficult to discern their appropriacy as an example of an item in a course of minimal aims. However, as Ellis fails to give examples of his simple formulae, it is possible that this text does not contain examples thereof. If this is so, one wonders why the text is cited at all.

<sup>3</sup>Stevick 1980, p. 289) states in this regard, “There remains, I am afraid, a residue—not universal but widespread—a residue of resistance and resentment against being given opportunities instead of rules and vocabulary lists . . . for which one foresees a practical need.” This is compatible with the findings of Willing (1988) in his research on Australian immigrants. I suspect that Japanese students fall into this category, thus adding an additional burden on CR by means of problem solving.

<sup>4</sup>It is somewhat surprising that Ellis makes no mention of other research related to PS for, although there are no studies which demonstrate the superiority of a PS methodology over others, Winitz and Read (1975) report extensive research into the efficacy of PS. Unfortunately, although they make strong claims for it (*ibid*, p. 24), they provide no substantial comparative data to support their position. In fairness to Ellis here, it should be pointed out that these two authors conceive of PS somewhat differently to him. They view it more in terms of the hypothesis testing of the L1 learner.

As one might expect, PS has been the subject of substantial research in the field of psychology (Ernst & Newell, 1969; Miyake, 1986). However, as Ellis makes no reference to research within his own field, it is hardly remarkable that studies beyond it remain undisturbed.

### References

- Ellis, R. (1991). Communicative competence and the Japanese learner. *JALT Journal*, 13, 103-130.
- Ernst, G. W., & Newell, A. (1969). *GPS: A case study in problem solving*. New York: Academic Press.
- Miyake, N. (1986). Constructive interaction and the iterative process of understanding. *Cognitive Science*, 10.
- Rutherford, W., & Sharwood-Smith, M. (Eds.). (1988). *Grammar and second language teaching*. New York: Newbury House.
- Sharwood-Smith, M. (1981). Consciousness raising and the second language learner. *Applied Linguistics*, 2, 159-168.
- Sheen, R. (1990). The applied linguistics gap. *English Teaching Forum* 27 (3), 47-48.

- Stevick, E.W. (1980). *A way and ways*. Cambridge, MA: Newbury House.
- White, L., Spada, P., Lightbown, P. M., & Ranta, L. (1991). Input enhancement and L2 question formation. *Applied Linguistics*, 12, 416-432.
- Willing, K. (1988) *Learning styles in adult migrant education*. Adelaide: National Curriculum Resource Centre.
- Winitz, H., & Read, J. (1975). *Comprehension and problem solving as strategies for language learning*. Mouton: The Hague.

**Ronald Sheen, Tottori University**

### **Rod Ellis Replies**

In writing "Communicative Competence and the Japanese Learner," I hoped to stimulate discussion of the role of communicative language teaching in high schools and colleges in Japan. I welcomed Sheen's response, therefore, in the expectancy that it would stimulate debate.

In a comprehensive survey of modern rhetoric, Tootell (1992) proposes a number of principles to guide the writer of a "critical response" of the kind Sheen provides. I will make use of these principles as a way of shaping my own reply.

**Principle 1: Make sure that you represent the author's views accurately; avoid the "straw man" error.**

Sheen claims that I have failed to take account of the "essential characteristics" of English classrooms in high schools in Japan (he makes no mention of colleges), which he says are characterised by "a stringent examination system demanding formal grammatical knowledge" and "teachers, students and parents . . . give priority to the principles of this system." In fact, though, I explicitly acknowledge this problem. On page 119 I refer to "the importance currently attached to grammar in public examinations and the widespread use of Japanese as the language of classroom communication." Sheen also states that I make the "implied claim . . . that a problem-solving activity is more effective than any other available method," but Sheen does not show in what way this "claim" is "implied." In fact, in this particular article I seek only to advance arguments in favour of problem-solving activities and carefully avoid comparing them with other activity types. It would appear that Sheen has fallen into the "straw man" error.

## Point-to-Point

### **Principle 2: Deal with the principal arguments advanced by the writer; avoid the “sideshow” error.**

One of the main purposes of my article is to demonstrate that in order to give serious consideration to how to teach communicative competence it is necessary to make a clear distinction between “knowledge” and “control” on the one hand, and between “linguistic” and “functional” knowledge on the other. The discussion of these distinctions is the core of the article—it covers seventeen out of twenty-five pages of text—and provides the theoretical basis for the “modest” (Sheen’s word) pedagogic proposals which I subsequently advance. Sheen fails even to mention the theoretical framework I develop, let alone discuss it. It would appear, therefore, that he is guilty of the “sideshow” error.

### **Principle 3: Present clear and full argumentation of any central thesis; avoid the error of an unsubstantiated assertion.**

One of Sheen’s major objections to my proposal regarding the use of problem-solving activities for developing learners’ sociolinguistic knowledge is that I do not provide any “empirical support.” He goes on to state that applied linguists “should limit discussion of . . . ideas for major changes in methodology to an audience of fellow applied linguists until such time as there is empirical support.” Here, clearly, is a very strong and controversial assertion. If it had been adhered to, for instance, it would have precluded the publication and implementation of Widdowson’s ideas for language teaching, as Widdowson (a noted applied linguist) has been singularly lacking in the provision of any “empirical support” for them. I scanned Sheen’s response carefully for any justification of this extraordinary claim, but could find none. It would seem that he is guilty of the “unsubstantiated assertion” error.

### **Principle 4: Offer a constructive alternative to the ideas you are seeking to refute; avoid appearing in a totally negative light.**

My “modest proposal” was that (a) an attempt should be made to

## Point-to-Point

teach sociolinguistic knowledge (but not sociolinguistic control), and (b) this might be achieved by the use of problem-solving activities. Sheen does not consider (a) and rejects (b). But what are Sheen's own views? Should we teach sociolinguistic knowledge? And, if so, how can it best be done? Unfortunately, we are only told what Sheen does not like in my proposal; there is no constructive alternative.

### **Principle 5: Make you points in a neutral and objective manner; avoid being polemical.**

Sheen has a taste for the polemical. He talks about my "intuitively seductive argument" (a metaphor that I am not at all happy about!). He talks about me "not heeding advice" from Rutherford and Sharwood-Smith, and my proposal being "bereft of empirical support." Emotive stuff. He also offers some flattery to placate me, however: I am an "applied linguist of deserved reputation and influence."

I have tried to demonstrate how Sheen fails to adhere to Tootell's rhetorical principles for the writing of critical responses. My purpose, however, is not so much to criticize Sheen's rhetorical skills as to point out to the reader the difficulty of trying to engage in serious debate on the basis of his response.

What I would have liked to have seen is some discussion of my proposal for a "minimal goal of language education in Japan" (i.e., teaching "knowledge" but not "control"). This is controversial, as many teachers probably feel that some attempt should be made to teach "control." It is a proposal that needs debate. It is a pity that Sheen does not provide it.

### Reference

Tootell, A. R. I. S. (1992). *An advanced rhetoric for beginners*. Athens: Plato Press.

**Rod Ellis, Temple University Japan**

## Reviews

**AN INTRODUCTION TO THE PRONUNCIATION OF ENGLISH.** A. C. Gimson. Fourth edition, revised by Sarah Ramsaran. London: Edward Arnold, 1989. 364 pp. + xix.

**ENGLISH PHONETICS AND PHONOLOGY: A PRACTICAL COURSE.** Peter Roach. Second edition. Cambridge: Cambridge University Press, 1991. 262 pp. + x.

In a recent special issue of the *TESOL Quarterly*, invited pronunciation specialist Joan Morley presents a list of "excellent English language reference books" (1991, p. 488) that does not include either book reviewed here. According to a recent review in *Language*, however, "one can hardly conceive of a more comprehensive or better organized reference work" (Sietsema, 1991, p. 652) than Gimson. And according to Gimson's former collaborator, Susan Ramsaran, who revised and published the fourth edition for Gimson posthumously, "despite the increasing number of books concerning phonetics and phonology, none has been published to supersede this volume which has been the most comprehensive and authoritative work on the pronunciation of English for almost thirty years" (p. vii). Even Roach refers to Gimson as a "de facto standard" (p. 6).

Of these two revised works by Gimson and Roach, however, Roach's (first edition, 1983) is the better written and edited, and the more revised and current. Gimson's fourth edition includes only 14 (of 325) bibliographical entries that are not by Gimson or Ramsaran, and that are later than 1980, the year of publication of Gimson's third edition. Furthermore, although Gimson's fourth edition is 70 pages longer than the first edition, 27 of those pages comprise a section entitled "Teaching the Pronunciation of English" (pp. 312-339) that is confusingly referred to as both "Chapter 12" and the "Appendix"—apparently reflecting some uncertainty about what to do with the added section.

Ramsaran preserves a Gimsonian tone and style that reflects atti-

tudes and conventions of an earlier era. The fourth edition, as did the first, begins: "One of the chief characteristics of the human being is his ability to communicate to his fellows" (p. 3). The fourth edition still warns the reader that "a disturbing development concerns the use of English in the Indian and African continents," and prescribes action: "It is for this reason of universal importance that efforts should be made to relate these developing forms of English pronunciation to either a British or American model" (p. 91). But Gimson has passed away, Ramsaran (his former student) is only being faithful, and Gimson's work, one hopes, can be appreciated as a classic.

Roach's work, on the other hand, was conceived in the dawn of "world Englishes." Roach wastes no time in recognizing dialect equality and arguing for a British model of English on utilitarian grounds. Both Gimson and Roach describe the same model, "RP" (explained below), rather than "General British" or "International English." Of the two authors, Gimson devotes a greater proportion of his description to phonetics and segmentals; Roach, to phonology and suprasegmentals. Gimson's work is more than an "introduction"; it is a comprehensive description and reference for RP. Roach's work is more than "a practical approach" to English phonetics and phonology; his purpose is to explain the pronunciation of English in the context of a general theory (i.e., phonetics and phonology) about speech sounds. Roach includes "Written exercises," "Notes for Teachers," and "Notes on problems and further reading." Roach, unlike Gimson, appeals directly to teacher trainers and non-native English users in EFL situations, and includes two cassettes (not treated in this review) with script, exercises, and answers in the book's appendix.

**Treatment of basic concepts and segmentals.** Gimson's Part I, "Speech and Language" (pp. 3-58) introduces basic concepts: communication and speech; articulatory, acoustic, and auditory phonetics; phoneme, allophone, and complementary distribution; vowel and consonant; the phonetic alphabet. Roach also devotes the first part of his text to basic concepts (pp. 3-66), but at the same time integrates his description of many English vowels and consonants directly into the

## Reviews

same discussion. Gimson describes the English sounds separately and in much more detail in his next section.

Gimson's Part II, "The Sounds of English" (pp. 59-219), forms the central part of his book. It begins with background information on English sound change, studies of pronunciation, and the model of English that he describes—"Received Pronunciation" or "RP." Although recent in origin (despite Roach's claim that RP "has always been chosen by British teachers to teach to foreign learners" [p. 5]), RP is the established accent of the educated English middle class, and the variety most promoted overseas by the British Council. British speakers may approximate or diverge from RP according to their feelings about the "Establishment." Roach describes RP as "the accent used by most announcers and newsreaders on *serious* national and international BBC broadcasting channels. It has for a long time been identified by the rather *quaint* name Received Pronunciation" (p. 4, emphasis added).

Gimson (but not Roach) distinguishes three types of RP: "conservative," spoken by the older generation; "general," used by the BBC; and "advanced," spoken by the younger generation. Gimson warns that "advanced RP" may be regarded as "affected" (pp. 88, 107) by speakers of the "general" and "conservative" varieties, just as RP (of any type) may be regarded as affected by speakers of other varieties of English (p. 88). Both Gimson (pp. 94, 315) and Roach (p. 24) describe RP as evolving and describe ongoing phonetic and phonemic changes.

According to Gimson, of the 300 million native speakers of English "only a minute proportion" speak RP (p. 90). Roach makes no claim in this regard. Others (e.g., Trudgill & Hannah, 1985, p. 2) have made more precise estimates: Only three to five percent of the speakers in England use RP. Gimson argues for teaching an RP model because it is "widely and readily understood," "it is adequately described in textbooks," and it "has ample recorded material available for the learner" (p. 315). Neither Gimson nor Roach addresses the fact that few native-speaking English teachers speak RP as a native dialect.

Gimson's description focuses on "general RP." This description,

forming the core of the book, has earned Gimson its many favorable reviews through four editions. He treats each phoneme (sound) consistently in five steps: (a) examples of spelling forms; (b) articulatory descriptions; (c) regional and social variants; (d) historical sources; and (e) advice to the foreign learner. The first three steps have earned what Sietsema (1991, p. 652) and Ramsaran (in Gimson, 1989, p. vii) must have been referring to in their praise. The fourth would be appreciated better in an uninterrupted history of the English sound system.

**Treatment of the teaching of pronunciation.** Why Gimson's fifth step ("Advice to the Foreign Learner") and the related pedagogical section toward the end of the book continue to receive favorable reviews (e.g., Sietsema, 1991) and recommendations (Roach, p. 7) is unclear. For "dark *l*," Gimson's advice is "hands on": "The tongue-tip may be gripped between the teeth during practice" (p. 206). For retroflex *r* the advice is a bit more abstract: "Once the feeling of slight curling-back of the tip and hollowing of the centre of the tongue has been achieved, the student should hold this position" (p. 211). Roach's description (not paraded as "advice") of retroflex *r* (and other sounds) is clearer: "If you pronounce an alternating sequence of *d* and *r* (*drdrdrdrdr*) while looking in a mirror you should be able to see more of the underside of the tongue in the *r* than in the *d* where the tongue tip is not raised and the tongue is not curled back" (p. 60).

Among the major changes in Roach's second edition is the elimination of any claims or attempts to deal with pronunciation teaching methods. This makes his book more focused, coherent, and faithful to its claims. In most of Gimson's sections on teaching pronunciation, people of different language backgrounds are lumped together as "the foreign learner" and advised in mass. This practice contradicts Gimson's (outdated) principle that "teaching should obviously be concentrated on those features of English which are not found in the learner's native language" (p. 318).

When Gimson does attempt to deal with native languages of RP learners (e.g., Polish, French, and Urdu), it is not at all clear what the

## Reviews

criteria are for selection of these languages. The index does not include languages or dialects—a serious limitation if this book is, as Ramsaran says (p. vii), a “reference” book. native speakers of other languages (e.g., Japanese) and other English dialects (e.g., American English) are, unfortunately, not able to turn directly to the special problems that RP poses for them.

**Treatment of suprasegmentals.** Both Gimson and Roach discuss suprasegmentals after segmentals. In Part III, “The Word and Connected Speech” (pp. 223-311), Gimson addresses “accent” (what Roach and others prefer to call “stress”), phonotactics (rules for the ordering of sounds), intonation, and phonetic variation in “connected speech” (defined as “an utterance consisting of more than one word” [p. 260]). Although Gimson states that “for all learners accentuation must provide the foundation on which any pronunciation course is built” (p. 318), he does not explain why this section follows the segmentals or why his “advice” is extremely brief.

Of the status of English as a “stress-timed” language, and the difficulty this poses for the adult learner, Gimson makes the unverifiable claim that English has “no exact parallel in any other language” (p. 318). Roach opts for a more defensible position: “Some languages (e.g., Russian and Arabic) have stress timed rhythm similar to that of English, others (such as French, Telugu and Yoruba) have a different rhythmical structure” (p. 121). Roach is more likely, here and elsewhere, to admit uncertainty: “English rhythm is a controversial subject” (p. 130).

**Recommendation.** With regard to their treatment of the teaching of pronunciation, the recent editions of Roach and Gimson have made very different choices. Gimson (through Ramsaran) has increased, rather than decreased, pedagogical advice that was often unsound, that never really belonged, and that distracted the reader from Gimson’s real value. Gimson’s book is to be recommended for what made it a classic: Insofar as it describes RP, it is, as Ramsaran claims, one of the most comprehensive and authoritative reference works available. But classics are to be reprinted, not revised.

In Roach's revised edition, on the other hand, the author has eliminated the pedagogical component of pronunciation. He has made his book a more focused and coherent textbook for the study of RP pronunciation, phonetics, and phonology. Although not as detailed as Gimson, Roach is clearer and more current. His "Additional Notes" are an excellent guide for those who wish to look further. For non-native English speakers, Roach's work may be the most straightforward and practical book on English phonetics and phonology to date. Budding authors of practical textbooks describing the pronunciation, phonetics, and phonology of RP or "any other English" should take note. Roach's second edition is the one they will have to emulate.

### References

- Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, 25, 481-520.
- Sietsema, B. (1991). Review of *An introduction to the pronunciation of English* by A. C. Gimson. *Language*, 67, 652-653.
- Trudgill, P., & Hannah, J. (1985). *International English: A guide to varieties of standard English*. Second edition. London: Edward Arnold.

**Reviewed by Tim Riney, International Christian University.**

## Reviews

**CONDITIONS FOR SECOND LANGUAGE LEARNING.** Bernard Spolsky. Oxford: Oxford University Press, 1989. 272 pp.

*Conditions for Second Language Learning* plugs an enormous gap in the literature on second language teaching and learning. It provides language teaching professionals with a theoretical model which simply and clearly sets out the multitude of interrelated conditions under which language learning or acquisition occurs. Previous theories of second language learning, as Spolsky points out, had more in common with advertising or promotional literature than with reasoned and logical scientific discourse (p. 2). Theorists were more interested in advancing the cause of their methodological biases than in the processes involved in language learning. As Spolsky notes, they disregarded the fact that "there are many ways to learn languages and many ways to teach them, that some ways work with some students and fail with others" (p. 15). Spolsky's theory is different. It is not linked to any method or way of teaching. It simply attempts to delineate the conditions under which second languages are learned or acquired and to explain why some learners are successful and others are not.

*Conditions for Second Language Learning* contains an introduction, 14 chapters, and an appendix. Chapter 1 outlines the characteristics of a general theory of language learning. The next four chapters deal with what knowing a language entails and the testing and/or measurement of language knowledge and/or ability. Chapters 6 and 7 discuss the effects of psycholinguistic and personal factors on second language learning. The next two chapters detail the linguistic basis and social context of second language learning. Chapters 10 through 12 discuss the effects of attitudes and motivation, opportunities for learning, and formal instruction on second language learning. The last two chapters focus on how to test the theory and what a general theory should be able to do. The appendix describes a case study.

Each chapter consists of three basic parts. The first provides the rationale for the inclusion of the chapter. The second reviews the literature relevant to the chapter. The third part relates the literature to the conditions.

In order to truly appreciate *Conditions for Second Language Learning*, it is necessary to understand what the nature of a general theory is for Spolsky. Among other things, a general theory must account for the complexity of conditions under which second languages are learned. It must also account for both success and failure in language learning. It must account for individual differences in language learning. It must account for some peoples' ability to learn more than one second language. It must be consistent with language acquisition theory. It must encompass all the various ways of learning languages. And it must be capable of being verified empirically. Spolsky's theory, unlike the theories he finds fault with (in his Introduction, Spolsky discusses the limitations and weaknesses of earlier theories of second language learning), possesses all of these characteristics. That it has these characteristics is due to Spolsky's decision to express his theory through conditions rather than to base it on a single overriding principle or to associate it with any particular teaching methodology.

*Conditions for Second Language Learning* is constructed around Spolsky's belief that the conditions which impinge on second language learning are of three distinct types. There are necessary conditions—conditions which must be present for second language learning to occur. An example is Condition #1: "A second language learner's knowledge of a second language forms a systematic whole" (p. 16). There are typical conditions, conditions which are commonly present, but which need not be. An example is Conditions #23: "The younger one starts to learn a second language, the better chance one has to develop native-like pronunciation" (p. 19). And there are graded conditions—conditions in which the degree to which a condition is satisfied influences the amount of language learning. Condition #23 is an example. A graded condition is unlike a necessary and typical one in that it cannot stand alone; it is always coupled to a necessary or typical condition. Thus, a condition can be necessary and either graded or ungraded, or typical and either graded or ungraded.

Many of Spolsky's 74 conditions are not all that original. Most have

## Reviews

been part of second language lore for years. Condition #1, for example, is an integral component of all interlanguage theories. And Condition #23 is something virtually all of humanity knows. Other conditions seem trite and self-evident. Examples of such conditions are #6: "When one learns a second language, one learns one or more varieties of that language"; and #8: "Individual language learners vary in their productive and receptive skills" (p. 17). that some conditions are unoriginal, trite, and self-evident does not demean Spolsky's contribution; any comprehensive theory must incorporate much that is widely known or accepted.

More interesting than the actual conditions is that Spolsky's typology of conditions make explicit what intelligent and observant language teachers have long based their teaching on. Teachers try to create in their classrooms the conditions necessary for language learning. They also encourage their students to do what they have observed to be typical of successful language learners. And they are aware that how well one learns a language is dependent on a host of gradeable factors: for example, learners who study more usually learn more. That the typology of conditions reflects or relates to what teachers know and do is not accidental. Spolsky believes that a general theory must have a sound pedagogical basis. Thus, there is much in *Conditions for Second Language Learning* which should benefit, interest, and appeal to most readers.

Spolsky's review of the research on language learning and acquisition should certainly interest most teachers. For language teachers, foreign language ones in particular, keeping abreast of developments in second language acquisition can be difficult. They need more books, like *Conditions for Second Language Learning*, which concisely and cogently summarize all but the most recent developments. While much of the research Spolsky discusses will be familiar to teachers trained during the last decade, some will be unfamiliar enough to cause teachers to alter some of their ideas about language learning. This reviewer certainly altered some of his.

Spolsky's review of the literature on the amount of time one must study a language in order to become fluent is revealing. He reports that French immersion students in Canada require 4,500 hours of instruction to attain high levels of fluency (p. 190). One shudders to contemplate the number of hours Japanese students require to reach similar levels of fluency in English. The number must be so astronomical that one begins to question whether the much maligned Japanese teachers of English are doing as badly as is commonly believed.

Much of the other research Spolsky reviews may also lead readers to alter their views. This reviewer found Spolsky's review of the research on the influence of personality and attitudinal variables particularly enlightening. The research he cites indicates, for example, that motivation is not a critical factor in short term language learning success. It may, however, be an important factor over the long term. More highly motivated learners seem to continue their studies for longer periods of time. And, curiously, it does not seem to make much difference whether one is instrumentally or integratively motivated (Ch. 10). And while teachers may believe that extroverts are better language learners than introverts, the research does not support this belief (p. 112). Even aptitude may not be that critical a factor in language learning success (p. 108).

This excellent book (the British Association for Applied Linguistics awarded it its 1990 Book Prize, and the Modern Language Association of America awarded it the tenth Kenneth W. Mildener Prize) has few, if any, detracting flaws. That it has few flaws is no doubt due to the refreshingly evenhandedness of Spolsky's approach. He examines the facts as they are presently known and does not manipulate them to support any preconceived notions he might have had. He simply wants to advance the process towards a coherent and definitive theory of second language learning, a process and a goal he obviously views as important (Spolsky, 1990).

While reading *Conditions for Second Language Learning*, readers must constantly keep one thing in mind. As the last sentence of the

## Reviews

previous paragraph implies, Spolsky has not written the definitive theory of second language learning. He has proposed a model which may or may not develop into a definitive theory. Whether it does will depend on future research and on whether it accounts for second language learning better than competing theories (see *TESOL Quarterly*, Winter 1990, for a review of current opinions).

Although this book is a worthy and welcome addition to the growing body of second language literature, it is not a book for everyone. To support his theory, Spolsky draws upon an enormous body of literature (the book is accompanied by a 21-page bibliography) from a wide variety of fields. Without some familiarity with psychology, linguistics, neurolinguistics, statistics, language acquisition, second language theory, transformational grammar, etc., readers may find many of Spolsky's arguments difficult to follow. Nevertheless, anyone who does read through the book—accepting that the going may not always be easy—will reap a bountiful harvest. This reviewer certainly did. *Conditions for Second Language Learning* is a book which he will consult regularly in the future.

**Reviewed by Richard J. Marshall, Toyohashi University of Technology.**

## References

- Spolsky, B. (1990). Introduction to a colloquium: The scope and form of a theory of second language learning. *TESOL Quarterly*, 24, 609-616.
- TESOL Quarterly*. (Winter, 1990). Vol. 24, No. 4. This entire issue is devoted to second language learning theory. Edited by Spolsky, it contains five articles by noted scholars on various aspects of second language learning theory.



# The Modern Language Journal

Founded: 1916

Editor: David P. Benseler  
Dept. of German  
Ohio State University  
Columbus, OH 43210

... is recognized throughout the world as the outstanding foreign language/pedagogical research journal in the United States. This popular journal was founded in 1916 by the National Federation of Modern Language Teachers Associations. Topics covered in MLJ include teaching strategies, bilingualism, applied linguistics, teaching of literature, study abroad, innovative foreign language programs, and developments in curriculum, teaching materials, and testing and evaluation.

ISSN 0026-7902

**Published quarterly at The University of Wisconsin Press**

Subscribe now, or recommend a subscription to your library.  
A detailed brochure will be sent upon request.

#### RATES

Individuals: \$17.50/yr.  
(must prepay)  
Institutions: \$35/yr.  
Foreign subscribers add \$8/yr.  
for regular postage, or \$20/yr.  
for Airmail delivery.

#### REPLY TO

Journal Division  
The University of Wisconsin  
Press  
114 North Murray Street  
Madison, Wisconsin 53715  
USA  
(608) 262-4952

---

**SECOND LANGUAGE WRITING: RESEARCH INSIGHTS FOR THE CLASSROOM.** Barbara Kroll (Ed.). Cambridge: Cambridge University Press, 1990. 256 pp.

Barbara Kroll's book, *Second Language Writing: Research Insights for the Classroom*, is intended for future and for practicing teachers and researchers in second language writing. Her experiences as a teacher-trainer working without a textbook that reflected changes in the field in recent years led her to undertake the writing of this volume in collaboration with fifteen other well-known authors. Each of the thirteen chapters was written specifically for this book. Although the authors wrote separate chapters, they had access to the other chapters. As a result, the chapters refer to each other, with comments included on similarities and differences. Since the authors interact with each other, the book is more than a mere collection of articles.

The book has two sections, each briefly introduced and summarized by Kroll. Section I, *Philosophical Underpinnings of Second Language Writing Instruction*, consists of six chapters. The first, by Tony Silva, is historical, and presents four major post-war approaches. He gives criteria for an effective approach to second language writing, and lays the foundation for the remainder of the book. Chapter 2, by Ann M. Johns, discusses three types of L1 theory: process, interactive, and social constructivist. Chapter 3, by Alexandra Rowe Krapels, gives "an overview of second language writing process research." Ilona Leki discusses issues related to teacher response to student writing in Chapter 4. In Chapter 5, Liz Hamp-Lyons discusses the problems of assessing writing, covering both instruments and scoring, as well as other issues. The last chapter in this section, by Joan Carson Eisterhold, covers theories about the relationship between reading and writing in both L1 and L2.

Section II describes "a variety of specific studies, each focused on a different aspect of writing and/or the writing classroom, representing some kind of option for either the student writer or the teacher." Chapter 7, by Alexander Friedlander, considers the effects of L1 versus L2

planning on writing about different topics. In Chapter 8, Ulla Connor and Mary Farmer explore how teaching topical structure analysis (developed in studies of coherence) can help students revise their work. In Chapter 9, Kroll examines differences between essays written in class and at home to determine the effects of time limitations on composition. The next two chapters look at aspects of teacher response to student writing. In Chapter 10, Andrew D. Cohen and Marilda C. Cavalcante compare how teachers say they respond, and how their students say the teachers respond, with how the researchers observed the teachers responding. Chapter 11, by Ann K. Fathman and Elizabeth Whalley, looks at the effects on student writing of comments by teachers on form and content. Joy Reid, in Chapter 12, examines fluency and syntactic and lexical variables to see if different L1 groups writing on two topic types show quantitatively measurable differences. In the last chapter, Cherry Campbell looks at how language proficiency relates to the ability to integrate background reading into composition. Every chapter is followed by references. A sixteen-page index concludes the book.

The research consists of two types: that published elsewhere and summarized here, and that reported for the first time in this book. As is pointed out in a number of chapters, and particularly in Chapter 3, research in this area often lacks comparability, and when comparable, is often contradictory. This leads to the question of what kind of insights for the classroom are offered here. Perhaps the main insight is that the subject is far more complex and problematic than is generally admitted. Many of the chapters challenge us to compare our own philosophies with those of the different authors. Some, such as Chapters 4 and 5, raise fundamental issues which we must face and resolve every day in our classrooms. Research may give hints and raise new issues, but it has given us few answers as yet. And Chapter 1 points up the size of the gap between modern theory and the quality and availability of composition textbooks—a fact which those of us who teach composition are all too aware of.

In Section II, Chapter 9 points out the need for improvement in teaching composition in Japan. Of advanced learners from five L1s

## Reviews

(Arabic, Chinese, Japanese, Persian, and Spanish), the Japanese scored far below their classmates in both accuracy and rhetorical competence in both test conditions. This contradicts the frequently made claim (by Japanese) that Japanese write well, and challenges us to reconsider how we teach composition here in Japan. Clearly, there is both room and need for improvement. The chapter which perhaps comes closest to demonstrating one way to improve performance is Chapter 11, which addresses the effects of feedback on writing. Several other chapters offer stimulating ideas to try out, such as topical structure analysis (Chapter 8).

This volume offers an up-to-date overview of issues in second language writing. It also presents results of research related to several issues in the field. However, as is pointed out many times, the need for further research in the field is acute. For this reason, the present book is not definitive. It tells us where we are and suggests directions for the future. Anyone interested in research can find a number of valuable suggestions here for further study. People involved in the teaching of reading may also be interested because of the close relationship between reading and writing. Of course, any teacher of second language writing will be interested in the book.

**Reviewed by Sandra S. Ishikawa.**

# CROSS CURRENTS

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

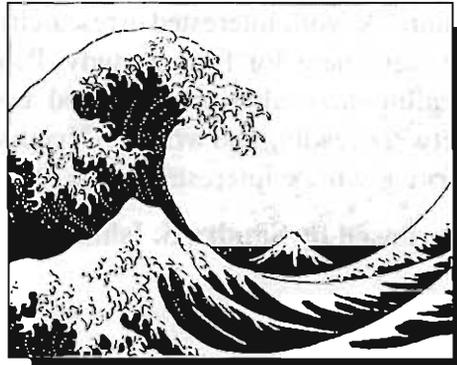
Editor: Thomas Clayton, The Language Institute of Japan, Odawara, Japan

*Cross Currents* is a biannual 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. Our goal at *Cross Currents* is to act as a medium through which international ESL/EFL issues, theories, and methodologies can be made available to our Japanese audience, and also through which Japanese ESL/EFL contributions—increasingly vibrant, innovative, and influential—can be disseminated to our international colleagues.

*Cross Currents* has served a diverse international readership continuously since 1972.

A Selection of Recent Articles and Essays:

- Alan Maley (UK)**, Courses or Resources  
**Subhash Jain (India)**, Arriving at the Himalayan Range  
**Thomas Clayton (Japan)**, Politics and the Modern English Language  
**John Staczek (USA)**, Professionalism and the Degree in TESOL  
**Robert O'Neill (UK)**, What Did Isadora Duncan and Einstein Have in Common?  
**John Swales (USA)**, Teacher-Researcher: Personal Reflections



SUBSCRIPTIONS	Single Issue	1 Year	2 Years
Inside Japan †	¥1,300	¥2,600	¥5,150
Outside Japan—Individual † *	US \$8.00	US \$15.00	US \$25.00
Outside Japan—Institution † *		US \$19.50	US \$35.00

† Special rates available for JALT members. Please use the postal order form in *The Language Teacher*.

\* Subscriptions may be ordered through your local bookseller or directly from *Cross Currents*, LIOJ, 4-14-1 Odawara, Kanagawa 250, Japan. Checks must be payable to *Cross Currents* (LIOJ), in US funds, and drawn on a US bank. For airmail, include an additional US \$3.00 for a single issue, US \$6.00 for a one-year subscription, and US \$12.00 for a two-year subscription. Back issues available through *Cross Currents*.

**CURRENTS OF CHANGE IN ENGLISH LANGUAGE TEACHING.** Richard Rossner and Ron Bolitho (Eds.). Oxford: Oxford University Press, 1990. 268 pp. ¥6,000.

This volume offers a selection of articles which appeared in the *English Language Teaching Journal* between 1982 and 1988. The intention of the editors, Richard Rossner and Ron Bolitho, is to provide readers with an overview of some of the major issues which have been in the forefront of developments in EFL/ESL. There are some twenty-three articles divided more or less equally into the following three sections: (a) the role and purpose of ELT; (b) the second decade of the communicative era; and (c) issues in methodology and teacher training.

**The role and purpose of ELT.** The first section addresses cultural aspects of ELT. The first two authors, J. Rogers ("The World for Sick Proper") and G. Abbot ("Should We Start Digging New Holes?"), discuss the advantages and disadvantages of the mass teaching of English in Asia and Africa. The former suggests that we limit such teaching to those who can best profit from it and not raise the hopes of those who cannot. G. Abbot, on the other hand, while sympathizing with some of Roger's arguments, rejects his suggestions on the grounds that we cannot predict who such students will be. C. and M. Alptekin ("The Question of Culture") and L. Promodou ("English as Cultural Action") deal with the issue of the cultural content of courses, both arguing, from different points of view, for the integration of the local variety into materials. H. Krasnick ("Images of ELT") raises an issue which should stimulate reflection in all teachers. He maintains that teachers tend to adopt a superior, counter-productive attitude toward students. He quotes striking examples to support this contention.

**The Communicative Era: Second Decade.** This section opens with a contribution from J. C. Richards ("Communicative Needs in Foreign Language Teaching"). It elaborates on a number of generalizations about communication, such as the fact that it is "meaning-based," "conventional," and "appropriate." Li Xiaoju ("In Defence of the Communicative Approach") describes the implementation of a com-

municative approach in China. He does so in a somewhat uncritical manner, polarizing methodological issues by emphasizing the defects of the grammar-translation and structural approaches. He would have strengthened his position had he produced results of this three-year project demonstrating the advantages of the communicative approach. This he conspicuously fails to do.

M. Swan's "A Critical Look at the Communicative Approach" might well be considered the outstanding contribution in this volume. It examines in detail the basic tenets of the approach and finds them wanting. Given the limitations of space, one example of this will have to suffice. Widdowson (1978) has pointed out that utterances vary in meaning according to variation in situation, and that this has crucial relevance for language learning. Swan quite rightly points out that this is a feature common to all languages and that, therefore, one need not devote time to explaining this to learners. Does one, for example, as a learner of Japanese need to be told that one might use the remark "*Samui desu ne?*" to hint that one might appreciate the turning up of the heat or the closing of a window, as well as simply remarking on the coldness of the weather? Surely not.

Widdowson ("Against Dogma: A Reply to Michael Swan") responds to Swan's criticisms. He does so either by intentionally or unintentionally misinterpreting Swan's arguments or by conspicuously failing to address the bulk of Swan's contentions (see Tredigo, 1986, for a similar reaction to Widdowson's contribution). Medgyes ("Queries from a Communicative Teacher") expresses the difficulties encountered by non-native speakers in attempting to implement communicative methodology. There will be many who will recognize the problems he describes.

In "Talking Shop," R. Allwright and Rossner interview Corder, who makes a number of points concerning current methodology, some of which appear to have little empirical support. Consequently, one would have expected the interviewers to be a little more probing in their questions. For example, Rossner inquires, "Are we any closer now to

## Reviews

understanding how learners learn?" "Very significantly so," Corder replies, offering as evidence the unconvincing findings of the morpheme order acquisition studies (see McLaughlin, 1987, pp. 31-34, for a review of research which casts serious doubts on the validity of the invariant morpheme acquisition hypothesis. Later in the interview, Corder maintains that teachers should concentrate on the teaching of vocabulary "on the grounds that grammar will look after itself." Substantial research evidence in support of this claim is lacking (see Long, 1988, for a review of the relevant research, but particularly of the effects of formal classroom language instruction). If interviewers allow to pass unchallenged such statements by leading figures in the field, they do a disservice to the readership. The other "Talking Shop" article in the third section, in which C. Gattegno is interviewed by Rossner, suffers from a similar defect. Gattegno makes various unsubstantiated statements which the interviewer allows to pass unchallenged.

**Issues in Methodology and Teacher Training.** Maley ("New Lamps for Old") gives a summary of developments of the decade before 1983. It will prove useful to those who are new to the literature. However, it is a largely uncritical review. Allwright ("What Do We Need Teaching Materials for?") and R. O'Neill ("Why Use Textbooks?") offer two diametrically opposed views on the value of commercially produced materials. Striking differences characterize the opposition. Allwright's is research-based and fully referenced. O'Neill's is experience-based and contains no references. However, the latter's arguments are more convincing, at least from this reviewer's point of view.

The field of EFL is characterized by numerous myths masquerading as received wisdom. One of these maintains that video will facilitate comprehension. I. Macmillan ("Video and Language Comprehension") throws doubt on this, showing that research findings indicate that visual information accompanying verbal information may cause reduction in comprehension. This is an important finding worthy of further research in ELT.

Wenden ("Helping Language Learners Think about Learning") provides teachers with a useful approach in analyzing students' learning styles and strategies. However, caution is advisable before embarking on such an approach. The issues involved here are complex, as is indicated by the extensive recent literature on the subject (see Oxford & Crookall, 1989, for a review of the relevant literature). The next article is by D. King ("Counselling for Teachers"). It is a short but informative paper on counselling and its relevance to teaching.

R. Nolasco and L. Arthur ("You Try Doing it with a Class of Forty") describe a project entailing the establishment of a communicative approach in Moroccan schools. In doing so they attempt to counter the objections of those teachers involved. They do so convincingly, providing that is, that one can be convinced by arguments without empirical support. E. Ramani ("Theorizing from the Classroom"), J. Harmer ("Balancing Activities"), G. Carter and H. Thomas ("Dear Brown Eyes"), and R. Ellis ("Activities and Procedures for Teachers") contribute articles related both to teacher training and teaching. Teachers and teacher-trainers will find these articles both useful and informative.

One might judge a collection such as this on the basis of the sum of the qualities of the individual papers. However, to do so would entail failing to evaluate the extent to which the editors have been successful in achieving their aims: that is, to provide a reflection of the field between 1982 and 1988 as shown in *ELT Journal* articles. During this period, *Journal* articles have dealt with a wide range of interests relevant to practicing teachers. This collection covers most of them. However, it is surprising to find no articles on testing or listening comprehension, two areas which have been both important issues in the field and the subject of papers published in this journal. It is also rather surprising to find that seven of the articles in the collection come from two consecutive issues (numbers 2 and 3 of volume 40, 1986). During the period covered by this volume, there were some twenty-eight issues of the journal. One would have expected a more even distribution.

In spite of these criticisms, the volume does provide an informative and sometimes stimulating account of most of the major issues. Practicing teachers will find it a useful resource, providing that is, as the

## Reviews

editors imply in their conclusion, they approach the book in a spirit of inquiry but with a dash of healthy skepticism.

**Reviewed by Ronald Sheen, Tottori University**

### References

- Long, W. (1988). Instructed interlanguage development. In L. M. Beebe (Ed.), *Issues in second language acquisition*. Rowley, MA: Newbury House.
- McLaughlin, B. (1987). *Theories of second language learning*. London: Edward Arnold.
- Oxford, R., & Crookall, D. (1989). Research on language learning strategies: Methods, findings and instructional issues. *Modern Language Journal*, 73.
- Tredigo, P. (1986). Correspondence. *ELTJ*, 40.
- Widdowson, H. G. (1978). *Teaching language as communication*. Oxford: Oxford University Press.

 Apple Computer  
Authorized Distributor

## GET YOUR HANDS ON A MACINTOSH!!

*Best* JALT-CATENA Education Program

Through the JALT-CATENA Education Program, any JALT member can purchase Apple Computer products at special prices from the ComputerLand shops specified by CATENA.

- 25% discount from the standard retail price for the LC and Classic II models.
- 30% discount for all models except the LC and Classic II.

Simple & Best, Macintosh!!

 Power Book

 Classic II

The world's largest personal computer retailer

# ComputerLand

CATENA CORP.

2-10-24, SHIOMI, KOTO-KU, TOKYO, 135  
For more information please call Toll Free 0120-332255

The Apple name and logo are registered trademarks of Apple Computer, Inc. Macintosh is a trademark of Apple Computer, Inc.

## Information for Contributors

### EDITORIAL POLICY

***Note: Submissions must conform to the Guidelines printed below.***

The *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 more than 20 pages, including reference list, typed on A4 or 8.5" X 11" paper, and double-spaced. **This is a strict requirement.** Writers must supply camera-ready diagrams or figures (if any) before final publication.

#### Materials to be Submitted

- Three copies of the manuscript
- Abstract (less than 200 words)
- Japanese translation of title and abstract if possible
- Running head title (one to four 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 your 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 the following address:

Roger Davies,  
*JALT Journal* Reviews Editor  
Nagasaki Prefectural University  
123 Kawashimo  
Sasebo-shi, Nagasaki 858  
(0956) 48-6088

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

Advertising  
JALT Central Office  
Shambaru Dai-2 Kawasaki #305  
1-3-17 Kaizuka, Kawasaki-ku  
Kawasaki-shi 210  
Tel: 044-145-9753  
Fax: 044-145-9754

---

JALT JOURNAL 第14巻 第1号

平成4年4月25日 印刷

平成4年5月1日 発行

発行人 全国語学教育学会 (JALT)

発行所 JALT本部

〒210 神奈川県川崎市川崎区貝塚1-3-17

シャンボール第2川崎305号

TEL (044) 245-9753

FAX (044) 245-9754

印刷所 株式会社 厚進社

〒530 大阪市北区天満2丁目13-3

TEL (06) 351-8795

---

