昭和五十四年四月二十一日 第三種郵便物認可
第十巻 第一・二号 昭和六十三年十二月一日発行

ISSN 0287-2420
¥750

JALT
JOURNAL

JOURNAL OF THE JAPAN ASSOCIATION
OF LANGUAGE TEACHERS

全国語学教育学会

VOLUME 10 NOS. 1 & 2
NOVEMBER 1988

FOR JIM WHITE

Hooray for EFL/ESL programs for all ages!

Learning fun
and teaching ease
every step of the way

- **My English Book Series**
  4 levels for children, ages 3-6
- **Hooray for English!** ***NEW***
  6 levels for children, ages 6-12
- **English for a Changing World**
  6 levels for children, ages 10-17
- **English in Tune**
  4 levels for teens and adults

For more information contact:

James Miller, Educational Consultant/Representative
6-12-4 Jingumae • Shibuya-ku • Tokyo 150, Japan
phone: (81-3) 499-4710
telex: J28899 SIBINITH ATTN; SIB620 MILLER

Scott, Foresman International Division
1900 East Lake Avenue, Glenview, IL 60025 Telephone: (312) 729-3000
Telex: 724371 FAX Number: (312) 729-8910 Cable: SCORESMAN, Glenview, IL U.S.A.

JALT Journal Volume 10, Nos. 1 & 2
November 1988

This issue is dedicated to Jim White, immediate past president of JALT, in gratitude for his years of work for JALT.

CONTENTS

Editorial 8
In This Issue 9

Articles

Abstracts in Japanese 10

Testing Listening Comprehension in Japanese University Entrance Examinations
Gary Buck 15

Yakudoku: Japan's Dominant Tradition in Foreign Language Learning
Nobuyuki Hino 45

Factors Influencing the Acquisition of Second Language Phonological Competence: Children versus Adults
Yasushi Sekiya 54

A Reassessment of Error Count
Jonathan Picken 79

To See Our Text as Others See It: Toward a Social Sense of Coherence
Daniel Horowitz 91

Point to Point

Paraphrase, Contemporary Poetry and Literary Syllabuses in Japan
Richard Cauldwell 101

Book Review

The Classroom and the Language Learner (Leo van Lier)
Reviewed by Tom Hinton 107

Information for Contributors 112

JALT Journal

Volume 10, Nos. 1 & 2
November 1988

Editors

Richard Cauldwell
Kobe University

Charles Wordell
Nanzan University

Book Review Editors

Rita Silver
Jane Wieman

Japanese Language Editor

Kyoko Nozaki

Advisory Board

Richard Berwick
Kobe University of Commerce
John Fanselow
Teachers College, Columbia University
John Fry
The British Council, Kyoto
Michael Horne
Nagoya University
Teruyuki Kume
Kobe University of Foreign Studies
Lawrence Schourup
Kobe University
Andrew Wright
Nagoya University

Executive Secretary, Thomas N. Robb
Publications Chairperson, Ann Chenoweth
The Language Teacher Editors, Eloise Pearson, Ann Chenoweth

National Officers 1988

President
Deborah Foreman-Takano

Vice-President
Tatsuya Komatsu

Treasurer
Philip Crompton

Recording Secretary
David Hough

Programs
John Laing

Membership
Sonia Yoshitake

Public Relations
Kazunori Nozawa

Chapters

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

EDITORIAL

Jim White

This issue of *JALT Journal* is dedicated to Jim White, immediate past president of JALT. We do this in gratitude for the immense contribution he has made in shaping the administrative structure and fostering the professional activities of JALT over the past decade. (We hasten to add that Jim is still very much alive.)

It was in 1978 that he first joined what was then the Osaka Chapter of JALT. Within a couple of months he had volunteered to help with refreshments at the chapter meetings. In 1980 he was Program Chairman, and by 1981 he was President of what had become the West Kansai Chapter. In 1982 he was elected President of JALT. He remained in that post until his resignation in August 1987.

One of his lasting contributions to JALT is seen in its administrative infrastructure. He personally created the handbooks for local chapter officers and the national treasurer; and he computerized and indexed the minutes of the Executive Committee meetings. In 1982, when his school was host to the JALT conference, he oversaw the computer registration, the administration of the display area, and the arrangements for student volunteers. He was also heavily involved with the production of the conference handbook.

He brought scholarly ability and technical skills to JALT having done research work in Educational Technology with the late Professor Nishimoto of International Christian University. Besides his JALT service he has always strongly supported the Language Laboratory Association. His expertise has earned him attention outside Japan. He was called in to advise on the feasibility study for the United States Information Agency's video project for teaching English as a Second Language.

As President, Jim was always willing to help in small as well as big ways. He kept and circulated audio and video tapes of lectures given by noted scholars and teachers. He was always willing to share information — about computer operation or chapter management — and was on call late into the night. For all this and much more, JALT remains forever in his debt. Thank you Jim.

IN THIS ISSUE

Gary Buck notes that English listening tests will be introduced into university entrance examinations, and sees a need for these tests to incorporate authentic language and tasks. This way, he argues, the washback effect on language classrooms and schools will be very positive. He reviews research on what the listening process involves, and then evaluates eleven types of listening test.

Nobuyuki Hino describes the historical background of foreign language teaching in Japan. The dominant tradition, "translation reading" has a history reaching back more than a thousand years. First used to teach Chinese, the tradition was continued in the instruction of Dutch and then English. The method is now deeply rooted in Japanese culture and is to many teachers the natural way of teaching languages. This article will be of interest to all who are interested in bringing about change in the field.

Yasushi Sekiya reviews the literature on the acquisition of native-like competence in pronunciation. In particular he compares the abilities of children and adults in reaching this goal. He finds that the received view that children, in the long run, are superior to adults is supported by the experimental evidence. He also notes that one commonly claimed reason for this difference, neurological lateralisation, is no longer thought to be a prime factor.

Jonathan Picken reports on a testing project in Holland in which *error-count* is used as a method of marking tests of writing. He argues that this method, though unfashionable, can be made to match the validity and reliability of other more fashionable ways of marking tests.

Daniel Horowitz urges that students, in learning to write, should focus on the intended readers and on their expectations of what a particular written genre should contain. He argues that this is a vital skill to be acquired by those who are learning to write; that this skill is encouraged by reading activities; and that reading and writing should therefore not be taught separately.

Also in this issue

Point to Point

Richard Cauldwell reacts to Alan Hirvela's article in the last issue of *JALT Journal* (9.2). He argues that teaching students to create paraphrases of contemporary poems is a useful aim for literature courses at university-level in Japan.

Book Reviews

Tom Hinton reviews a book on classroom research by Leo van Lier.

Notes

This issue of *JALT Journal* is Volume 10 nos 1 *and* 2. For some time the Journal has been publishing the two issues of each volume in different calendar years. This has created some confusion and a few difficulties, particularly for libraries. We wish to get our volume and issue numbers in step with the calendar. This is a technical change which will not affect the number of journals you receive in a year. The next issue of *JALT Journal* will be Volume 11 no 1, and will appear in May 1989.

The Editors would like to thank Charles Adamson, Malcolm Benson and Michael Liggett for help in the preparation of this issue.

The opinions expressed in *JALT Journal* are the authors' own. These opinions represent neither the policy of JALT nor the policy of the institutions in which they work.

HBJ, HRW and Cassell

HBJ —*Now Completed!*

Practical English 1, 2 and 3, Second Edition
by Tim Harris with Illustration by Allan Rowe
•Student Book •Writing Workbook •Instructors Manual •Cassette Tapes

HRW —*New Titles:*

Write It Down by Jeanne Portnoy-James
—A high-beginning to low-intermediate-level ESL composition text.

The Rinehart Handbook for Writers by Bonnie E. Carter and B. Skátes

The Rinehart Guide to Grammar & Usage
—Brief version of the Rinehart Handbook for Writers.

The Writer's Handbook by John Mc Kerman

Thinking on Paper by Judy Markline, Rose Brown & Robert Isaacson

—*New Editions:*

An Introduction to Language, Fourth Edition
by V.A. Fromkin and R.D. Rodman

A Glossary of Literary Terms, Fifth Edition
by M.H. Abrams

Cassell —*New Titles:*

Advanced International English
•Student Book •Teacher's Book •Cassettes

Cassell's Elementary Short Course by Martin Mulloy

Earth Sciences
General Engineering

HBJ
Harcourt Brace Jovanovich Japan, Inc.
Ichibancho Central Bldg. 22-1, Ichibancho, Chiyoda-ku, Tokyo 102
Tel: 03-234-3912 ELT Hotline: 03-710-5180 (Shari Berman)

ABSTRACTS IN JAPANESE

日本における大学入試のリスニングテスト

東京大学は、入試にリスニング（聴解）テストを取り入れることに決め、他の多くの大学でもこの東大の例にならうものと予想される。この小論文では、リスニングとは何であるか、またリスニングをテストする際の問題点、特に、日本における大学入試と関連あると思われる問題点を調べてみる。初めに、日本の大学入試の様式をみながら、テスト作成にあたっての限られた、または強要された条件のいくつかを考察してみる。次に、リスニングとは何であるかを調べ、いくつかの研究報告を検討しながら理論上最も重要だと思われる点を論議していく。これは教師及びテスト作成者が、リスニングテストを自ら評価出来るような基準をもたらしたいという試みからである。最後に、上で論議された基準に基づき、最もよく用いられているリスニングのテスト方法を検討してみる。特に、実際にテストを実施するにあたり、通常の授業への影響という観点から理論上の考察をする。そして、日本の大学入試に適するようなリスニングテストの作成にあたっての実用的提案を挙げる。(著者は大坂明浄女子短期大学助教授。又ランカスター大学博士課程で研究中。)

日本の英語教授法の主流は「訳読」、つまり、英文をまず逐語的に日本語に置き換え、それを日本語の語順に並びかえるという方法である。日本では、訳読技術の修得それ自体が英語学習の目標であると見なされる傾向が強い。この訳読は、漢文の訓読以来、一千年以上にもわたる長い歴史を有しており、日本の社会言語学的伝統の重要な部分を形成している。本稿では、日本人の外国語

学習における訳読の継承の過程を分析し、さらに訳読の伝統が今日の英語教育にどのような示唆を与えるかを考察する。

『第二言語の音声習得能力に影響を与える要因：子供と大人の対比において』

本稿では子供と大人の第二言語音声習得を比べる事により、音声習得能力に影響を及ぼすと考えられる要因を調べた。これまでの研究で、第2言語環境に於ける音声習得では子供の方が大人より優れているという事が分かっている。この原因としては大脳生理学説、習慣形成説、社会心理学説、入力説が考えられる。現段階では脳の側部化を音声習得能力低下に結び付ける大脳生理学説には無理がある。他3説に関しては、1つの説が決定的要因というよりも、全てが音声習得に関わっている可能性が大きい。これら3つの要因を考慮すれば、第2言語の音声習得能力に於ける子供の優位性のみならず、大人間の個人差がある程度説明出来る。

エラー・カウントの再評価

EFL writing における error-count evaluation の研究の歴史を振り返ると、この方法が analytic marking や general impression marking にとってかわる程充分発達した代案になるほど、研究されてきたとはとうてい言えない。

本稿では error-count evaluation の研究がこのような歴史をたどったのは、error-count evaluation そのものに本質的な欠陥があるのではなく、この方法に理論的基礎がないからであるということを強張する。この点をはっきりさせるため、オランダの CITO writing proficiency test(CITO,1984)と、それに関連した Melse and Verstralen(1986)に言及する。又、後者のデータを使って error-count evaluation を採用すべきであるという事を主張したい。

読者中心のテキスト：一貫性への追求

本稿では、読み、書きの基礎となる認知過程に関する文献を検討し、読み、書きにおける本質的な共通点とは「書かれたものの解釈(interpretation of written text)」であることを明らかにする。次に、解釈(interpretation)の概念を考察し、読者の用いる解釈の基準を認識することが、読者に即した一貫性(cohereuce)ある文を書く能力に不可欠であると論述する。L1、L2のいずれの場合も、書き手はこの種の認識を多読(extensive reading)によって得ると論じる研究もあるが、本稿ではさらに、精読(intensive reading)と特定のジャンルの分析も、文章作成能力に重要な貢献をすることを明確にする。

RELC JOURNAL

A JOURNAL OF LANGUAGE TEACHING AND RESEARCH IN SOUTHEAST ASIA

The **RELC Journal: A Journal of Language Teaching and Research in Southeast Asia** is one of the professional publications of the Southeast Asian Ministers of Education Organization (SEAMEO) Regional Language Centre (RELC) which is located in Singapore. The main purpose of the journal is to disseminate information and ideas on theories, research, methods and materials related to language learning and teaching. Articles covering a wide range of topics in applied linguistics appear in its pages semiannually in June and December. Special features include information on current research projects in Southeast Asia, book reviews and review articles on topics related to language teaching. In its thirteen years of publication, the journal has reached a wide audience in eighty countries and has been rated as one of the leading journals in applied linguistics throughout the world. Scholars of recognized stature have used it as a means of presenting their findings and obtaining feedback on important issues in language and linguistics.

GUIDELINES

A companion publication, **Guidelines**, provides language teachers with practical and innovative techniques to enliven and enrich teaching in the language classroom. Each issue focuses on a particular theme.

The **Journal** and **Guidelines** are published twice a year and a subscription may begin with the June or December issue. The annual subscription is US$9.00 for each title. Individual copies may be purchased at US$6.00.

Please send your order to The Publications Officer
SEAMEO Regional Language Centre
RELC Building
30 Orange Grove Road
Singapore 1025
REPUBLIC OF SINGAPORE

Announcing a newly acquired journal for 1988

JOURNAL OF NEUROLINGUISTICS

Editor: **Fred C C Peng,**
International Christian University, Japan

The **Journal of Neurolinguistics** is an international, interdisciplinary journal for the free exchange of scholastic ideas on the study of the interface of the Language Sciences and Neurosciences (in particular, Neurology and Neuropsychiatry). It provides a forum for the enhancement of knowledge and study of brain function in language behavior, experience and evolution of language and the brain, and accepts contributions that deal with such areas of specialization as Agrammatism, Language Disorders in Aphasias, Dementia and Language Disorders, Epilepsy and Language Disorders, Language Impairment in Parkinsonian Dementia, Language Disorders in Schizophrenia, Dementia in Modern Society, Deep Dyslexia and Acquired Dyslexia, or Sign Language in the Brain; that is, manuscripts pertaining to all aspects of cognitive functions in the brain that are relatable to language (and sign language) behavior, be they concerned with normal behavior or specific to behavioral deficits, are welcome.

Subscription Information
1989: Volume 4 (2 issues)

Annual subscription (1989)	DM 290.00
Two-year rate (1989/90)	DM 551.00

ISSN 0911-6044

Deutsche Mark prices apply in Japan only and are subject to change without notice.

For more information please contact:
Mr Ben Kato, Pergamon Press Japan, 5th Floor, Matsuoka Central Building,
1-7-1 Nishishinjuku, Shinjuku-Ku, Tokyo 160.

SEND FOR A FREE SAMPLE COPY

PERGAMON PRESS INC.
Fairview Park, Elmsford,
New York 10523,
USA

PERGAMON PRESS plc
Headington Hill Hall,
Oxford OX3 0BW,
UK

JALT Journal, Vol. 10, Nos. 1 & 2

TESTING LISTENING COMPREHENSION IN JAPANESE UNIVERSITY ENTRANCE EXAMINATIONS

Gary Buck

Abstract

Several major Japanese universities have recently included or are shortly to include an English listening comprehension section in their entrance examinations. It may be expected that a large number of other colleges will follow suit. This paper examines the nature of listening comprehension, and the problems of testing it, with emphasis on those problems relevant to constructing university entrance examinations. Section 1 looks at the importance of college entrance examinations in Japan, and their effect upon classroom instruction. Section 2 examines the nature of listening comprehension, reviews theories and research, and provides criteria for creating and evaluating listening texts. In section 3, the most common methods of testing listening comprehension are discussed in terms of these criteria, and in terms of practical implementation and washback effect in the classroom. Practical recommendations are made for constructing comprehension tests suitable for Japanese university entrance examinations.

Recently a number of universities, including Tokyo University, have added an English listening test to their other entrance examinations. Many other colleges are likely to follow this lead and include listening tests in their own entrance examinations. However, listening comprehension is an extremely complex process, about which little is known, and thus the testing of listening comprehension presents considerable problems. The purpose of this paper is twofold. First, to examine some of the complex issues involved in understanding listening comprehension, how these relate to the construction of listening comprehension tests, and to offer some practical suggestions about the advantages and disadvantages of most of the common methods of testing it. Of course, there are no simple, instant solutions to such

Gary Buck is an assistant professor at Osaka Meijo Women's Junior College, and is currently working towards a doctorate in testing second language listening comprehension at Lancaster University.

complex problems. Rather, it is necessary to examine the issues and make the best decisions possible in the light of the available evidence. The second purpose of this paper is to try to stimulate and encourage practical and helpful discussion of these issues, among both language teachers and testers.

1. The English Entrance Examinations

Japan is a country in which the entrance examination reigns supreme. It is almost impossible to overstate the influence of these examinations on both the educational system as a whole, and the day-to-day content of classroom teaching. Their importance in the lives of young people is such that almost all future social and economic advancement is dependent on the results of these entrance examinations.

1.1 Constraints on Entrance Examinations

However, despite the immense importance of these examinations, there are a number of practical constraints on their production which considerably influence the nature of the tests produced. With the exception of the first part of the entrance examination to national universities, entrance examinations are usually made by each individual college, and colleges generally have two or three different examinations every year, which leaves very little time for development. The actual examination is usually made by a committee of teachers in each college, very often on an annually rotating basis. Due to the large number of candidates and the necessity (or custom) of completing all the marking in one long session the same or the next day, the tests must be quickly and easily scorable by a large group of teachers who may not have been involved in the test-making process. Security is a continual problem, considering the importance of the results, and it is generally felt to be impossible to pre-test items in order to ascertain whether they function as the item writer intended. Nor is it possible to use items which seemed to work well in past years, as most schools feel they must publish each year's entrance examination for the benefit of future applicants. A further important constraint, which should not be overlooked, is the feeling that the examination reflects on the prestige of the college making it. This means in practice that a test should look difficult, to give the impression that the college has a very high standard.[1]

1.2 The Nature of the Tests

Given these considerable constraints, the actual tests themselves have to be put together quickly by a committee of teachers with little specialized training in Educational Measurement as a separate academic discipline. In such a situation most teachers see little alternative but to try either to produce a different version of the tests made in previous years, or imitate the tests of more prestigious institutions.

There has developed something which may be easily recognized as a typical Japanese college entrance examination, which seems to be based on no clearly stated view of language or language acquisition. Furthermore, there is rarely any attempt to test language in use, and the items used are often of a type not currently recommended by language testing researchers (Heaton, 1975, p. 186; Buck, in press). Finally, these tests are seldom subjected to any standard statistical analyses or other validation procedures; it is generally assumed by both test makers and test users that tests made by experienced teachers are automatically good tests. According to Spolsky's classification, these tests are clearly in the "pre-scientific stage" of language testing (Spolsky 1975).

1.3 Effects on Classroom Teaching

There is a natural tendency for both teachers and students to tailor their classroom activities to the demands of the test, especially when the test is very important to the future of the students, and pass rates are used as a measure of teacher success. This influence of the test on the classroom (referred to as *washback* by language testers) is, of course, very important; this washback effect can be either beneficial or harmful.

Most educators would probably agree that the content of classroom instruction should be decided on the basis of clearly understood educational goals, and examinations should try to ascertain whether these goals have been achieved. When the examination does that, it forces students and teachers to concentrate on these goals, and the washback effect on the classroom is very beneficial. However, if the examinations are not testing these goals, students, who usually have their sights set on the examination, may pass the tests but totally fail to attain the basic goals set by educational planners. In such a case the washback effect is very negative indeed. It is perhaps here that Japanese entrance examinations are most destructive.

Many English teachers in Japan want their students to be able to use English as a means of communication. However, as long as entrance examinations consist mainly of such items as discrete-point grammar questions and translation, then students are obviously going to concentrate their efforts on learning to answer grammar questions and do translations, despite the wealth of evidence which suggests that such activities will not lead to successful language learning.[2] (Krashen, 1982; Savignon, 1983; Richards & Rogers, 1986). If we want our students to learn to communicate in English, then we must give them tests which require them to process communicative English.

There are probably many reasons why most Japanese high school graduates cannot use English for even the most basic purposes, despite receiving hundreds of hours of classroom instruction, but surely one of the most important is the washback effect of entrance examinations on the classroom.

1.4 Possibilities for the Future

The new trend towards including a listening comprehension section in entrance examinations seems to offer an opportunity to make examinations which test "real" English and will thus have a positive washback effect on the classroom. Ideally the new listening tests should fulfill two roles. First, they should provide an accurate and fair measure of the English listening ability of testees. Second, they should have a positive washback effect on the classroom, so that when students study for the test, they will automatically have to engage in activities which will lead to effective language acquisition. Although listening comprehension is an extremely complex process, it is the writer's opinion that both these aims can be attained with a fair degree of success if test makers are prepared to confront the considerable problems involved.

It is obviously impossible to measure anything without first establishing exactly what is being measured. Therefore, in order to come to some sort of understanding of how listening tests should be made, the most obvious necessity is first to examine the nature of listening comprehension itself. This will provide a theoretical basis for teachers and test makers to develop their own critical standards, and thus enable them to evaluate listening tests for themselves.

2. Listening Comprehension

Unfortunately, it is very difficult to come up with a simple definition of what listening is, and what listening comprehension means. Instead we have to consider research from various fields and try somehow to bring this together and arrive at some workable conclusion.

A number of different, but related, skills are involved in listening. The listener must take in a stream of sound and somehow convert this into meaning. To do that it is clearly necessary to have knowledge of the language. This knowledge obviously includes the phonological, morphemic, syntactic and semantic rules of the language, as well as cohesive devices, text types, etc., and the ability, in practice, to apply such knowledge rapidly and automatically. Researchers also emphasize the importance of pragmatic and general world knowledge in language comprehension. Naturally in such a complex area different theorists or researchers produce different classifications of these sub-skills. The key area of disagreement, however, concerns the relationship between these skills, and how they affect each other. There are two main lines of thought: 1) that language processing is only bottom-up, or 2) that it is both *top-down* and *bottom-up.*[1]

2.1 Bottom-up Processing

The bottom-up approach sees language comprehension as a process of passing through a number of consecutive stages, and in each stage one particular sub-skill is utilized, starting with the lowest level and moving up to higher levels of processing. First, at the lowest level, the acoustic input is decoded into phonemes, and then the information obtained is used to identify individual words, after which processing continues on to the next stage, the syntactic level. Only after the syntactic level is completed is the semantic content of the utterance extracted. Processing is thus seen as occurring on a number of different levels, starting from the bottom and working upwards through one level to the next, in serial order. The output of each level thus becomes the input for the next higher level. Of course, this is an oversimplification of a number of different, extremely complex, psycholinguistic theories (for a review, see Dirven & Oakeshott-Taylor, 1984, 1985). However, the important thing to note about the bottom-up approach to language comprehension is that the results of processing on a higher level are not available for use at lower levels. It is, as it were, a one-way street.

2.2 Top-down Processing

Advocates of the top-down approach to language processing generally accept the importance of lower level information, and also the fact that processing takes place on a number of different levels. However, they suggest that processing does not occur in a fixed order, from bottom to top, but rather, different types of processing may occur simultaneously, or higher level processing may take place before lower level processing. Thus, the results of higher levels of processing may be available to facilitate processing at lower levels. This means that they think it is quite possible in practice to understand the meaning of a word before decoding its sound. In almost all situations, linguistic and non-linguistic, our experience leads us to have expectations about what will happen next. These expectations, or hypotheses, may be very precise or rather general but nevertheless we almost always have some idea about what is likely to come next when we are listening. In such cases it is not necessary to utilize all the lower level information available to us, we can just take in enough to confirm or reject our hypotheses. Comprehension, then, becomes largely a process of hypothesis generation and testing. For example, if we hear the following uncompleted sentence:

[1] She picked up the gun, aimed and (Grosjean, 1980, p. 281)

we probably need very little acoustic information to understand that the final word will be *fired.* When we listen, we will hypothesize that the last word will be *fired* and then probably process only enough of the acoustic input to confirm our hypothesis, or, if we feel really sure, we may not even bother to listen to the last word at all. Similarly, when we part from a friend, we may hear a word of farewell, not so much by processing what he says, but because he is waving to us and saying something as he walks away.

Indeed, research has shown that our expectations are so strong that they can often lead us to "hear" things which were quite different from what was actually said. One cognitive psychologist, on the basis of his research results, concluded that, "contextual information is able to control the lexical interpretation assigned to a given acoustic-phonetic sensory analysis" (Marslen-Wilson, 1980, p. 49). Which means that what we hear is often a product of what we expect to hear. Bruce (1958), in a famous experiment, found that when subjects were asked to repeat a sentence heard against a background of noise, what they repeated was greatly influenced by what they had been informed was the topic of the sentence. When told that the sentence would deal with 'sport', the listener correctly repeated *I tell you that our team will win the cup next year.* However, when informed that the topic was 'health', the listener heard the same sentence as *I tell you that our team has been free from injury all this year,* and when told that the topic was 'weather' the listener heard *I tell you that I see the wind in the south next year.*

Over the last few years the evidence for the top-down nature of language processing has become stronger and stronger, such that now in applied linguistics it is normal to discuss comprehension in terms of top-down processing; that is, as an interactive process which simultaneously utilizes information from a number of different sources. Faerch and Kasper (1986, p. 264) suggest that comprehension is a process relying on three types of information: 1) linguistic and other communicative input; 2) the listener's linguistic and general world knowledge; and 3) information from the context and the conversation up to that point. This position has been facilitated to a great extent by research into reading. Reading and listening obviously share much in common, in that they are both different aspects of the process of language comprehension. While they do clearly differ in a number of crucial respects, it does seem reasonable for those interested in listening to look at research into reading, and consider how much of this is applicable to listening.

2.3 Schema Theory

There has been a considerable amount of research into reading and it would be impossible to review it all here. However, the trend over the last few years has been to see reading less and less as a data-driven process, and more and more as a hypothesis-driven process (Alderson & Urquhart, 1984; Smith, 1985; Devine, Carrell, & Eskey, 1987). We use our knowledge of language and the world to generate hypotheses about what we will read and then extract enough data from the printed page to confirm or reject these hypotheses. The reason we can read about something we are familiar with much faster than something which is very new to us is because our knowledge helps us to guess what is contained in the text, and we thus need to take far less from the text itself. We provide more of the information ourselves. The recognition of this fact led Goodman (1970) to describe reading as a "psycholinguistic guessing game." For each word in the language, or situation in which we find ourselves, there are a whole host of other things which are associated with it in our minds. These associations, or *schema*, form the basis of many of our expectations of what we will read (Rumelhart & Ortony, 1977; Rumelhart, 1980; Carrell, 1987). Not only that, but common situations in our lives tend to follow a reasonably predictable pattern. A visit to a restaurant usually involves things such as looking at the menu, ordering the meal, eating it and then paying for it. This pattern forms a sort of universal script for a restaurant "story" (Schank & Abelson, 1977).

These schema (scripts can be regarded as complex schema) are one of our ways of ordering our experience of the world. They seem to play a crucial role in the process of hypothesis generation which is so central to language comprehension. The context thus becomes an important part of the listening comprehension process, because it provides important clues to help us understand the meaning. Obviously, if we fail to provide a context, we rob students of important resources they would normally use in the listening comprehension process. In normal daily language use, we virtually never have to decode utterances by using only bottom-up information, devoid of the context which enables us to use top-down processing strategies. Therefore, if we ask students to decode short decontextualized sentences, we are not testing listening comprehension at all, but asking students to engage in a very unnatural activity which seems to be confined largely to the second language classroom.

There is as yet very little research into the effects of schema on listening, but there is every reason to suggest that they play just as important a part in listening as they do in reading. Unpublished research by the author has indicated that schema are tremendously important in listening. A number of students were asked to listen to a selection of different passages. While listening the tape was stopped at a number of points and the students were asked to retrospect on their listening processes. Results indicated that students found it much easier to comprehend passages which accorded well with their own knowledge, even when it seemed clear that the language used was well within their linguistic capabilities. All listeners reported creating images in their minds, which were far more detailed and complex than the descriptions contained in the texts they were listening to. When asked where these images came from, they indicated things they had learned at school, films they had seen, or books they had read. In other words, they came from their schema associated with the descriptions they heard. Listeners often had strong expectations about what would come next in the passage they were listening to, so strong in some cases that they prevented the listener from hearing what was actually said when it differed from their expectations.

The analysis of this data is still in progress, but one thing is already very clear. Although lower level acoustic and phonetic input is, of course, tremendously important to speech processing, comprehension of language in context utilizes a considerable amount of top-down information.

There are many examples in the data of top-down processing failing for some reason, resulting in a lack of comprehension, even though there appeared to be quite enough linguistic input for comprehension to take place. One striking case shows the influence of expectations. A student was listening to a story about a girl in Africa who asked a witch-doctor to solve a number of burglaries. The student had lived in the United States and was expected to understand the story with ease. However, she didn't know the word *witch-doctor*, so she made the sensible assumption that this was some sort of detective. Based on this wrong assumption, she naturally expected the witch-doctor to proceed like any normal detective. However, instead, he performed a number of magic actions to find out who the criminal was. The description of these events should have presented no difficulty to a student of her English ability, yet she totally failed to "hear" clues that should have made it quite clear to her that this was not a normal investigation, but a series of magic rituals. Based on her expectations, she "heard" her own story of a criminal investigation, quite different from the one she listened to, and remained convinced to the end that she had understood the passage quite well.

2.4 Pragmatics

There are other reasons why lack of context often inhibits the comprehension of language. This is because even if we have understood the basic semantic meaning of an utterance, we often need to know the context in order to understand what was meant. Leech (1983, p. 4) illustrates this by the different uses of the verb to *mean* in the following sentences:

[2] What does X mean?

[3] What do you mean by X?

The real meaning of what is said, the pragmatic meaning, is often very different from the surface meaning, especially in interactive conversation. To give a simple and often used example, when someone says:

[4] Can you pass the salt?

they are not usually asking about our ability to pass the salt at all, but are making request that we do so. Similarly,

[5] Do that again and I'll thump you.

has the appearance of a command to do something again, but as we all know, in actual fact it is an injunction not to do it again, or maybe it's even an invitation to a fight. Only the context will tell us which.

[6] I'd keep my mouth shut if I were you.

could vary in meaning along a continuum from being a piece of friendly advice all the way to being a murder threat. Of course, the intonation would probably help us understand to some extent, but to understand the pragmatic force of this fully we would need to know something about the participants, the relationship between them, the nature of the information the speaker is referring to, and the conversation that preceded this. The context in which an utterance is used not only provides the necessary conditions for engaging in top-down processing, but also provides a background against which the semantic value of the utterance can be interpreted.

2.5 Reasonable Interpretation

If the spoken message doesn't accord with our knowledge of the world, then we are going to have serious problems understanding and interpreting what we hear. Listeners' comprehension will vary depending on how well the spoken message they are listening to does, or does not, accord with their own knowledge, and hence, how much it agrees with or differs from their expectations. A talk on a subject about which we know nothing, all other things being equal, will be more difficult to understand, and further, a text which in some way violates or contradicts our expectations will be even more difficult to understand and could cause considerable confusion even though the language may not be particularly difficult.

Where high levels of new information in a listening passage make hypothesis generation more difficult, listeners will tend to be selective, resulting in different listeners extracting different things from the same text depending on how well the new information fits in with, or supplements, their own current state of knowledge. Also, different listeners often have different motives for listening, have different interests and different needs. Listeners will pay more attention to these features of a text which they think are more interesting or more relevant to their view of what is important to them in the passage. This is a very important point because "different listeners will reasonably extract different parts of the text as more 'salient' to them..., and so build their mental representations of 'what the text was about' around rather different structures" (Brown & Yule, 1983, p. 100).

The corollary of all this is that there is no such thing as one "correct interpretation" of a spoken text. Of course, in practice there are likely to be some interpretations which we would all consider so wide of the mark that we could call them wrong (for example with the witch-doctor story above), and some interpretations which are generally agreed to be better than others. But it must be understood that each listener has to construct a personal interpretation of what we heard, and what is comprehended or understood may well differ from listener to listener. It is therefore more appropriate to talk of "reasonable interpretations" rather than "correct interpretations."

3. Testing Listening Comprehension

Given the complexity of the listening process, it is easy to see that many tests which are called tests of listening comprehension are not really testing listening comprehension in the full sense at all. Indeed, Brown and Yule express the problem thus:

> ... we find existing approaches to the assessment of listening comprehension based on a very insecure theoretical notion of what "comprehension" means. It is by no means clear that a great deal of what is currently tested in listening comprehension tests is necessary, or relevant, to the process of understanding the communicative event which the student has listened to. (1983, p. 100)

It is indeed true that there is, as yet, no such thing as the ideal listening test. There are, however, a number of different types of listening test, each with its own advantages and disadvantages. The type of test preferred will depend largely on how individual test makers evaluate these advantages and disadvantages in the light of their own requirements, which is one reason why test users should try to develop a good theoretical understanding of listening comprehension. However, whichever testing method is chosen, it will be necessary to make compromises. For convenience, these different listening tests can be divided into two main approaches. They are the *process approach* and the *product approach.*

3.1 The Process Approach

The first of these, the process approach, attempts to identify the various sub-skills or processes used in listening and then tries to assess whether the testee has mastered these or not. This is not so easy, as we have already seen that the process of listening is complex and can vary according to the world knowledge and interests of the listener. However, we do know that both listeners and readers use their linguistic and world knowledge to generate hypotheses about the language they are comprehending. They do not need to get all the information (either graphic or acoustic) contained in a message in order to comprehend, because the language contains more information than is necessary. It is, in other words, highly *redundant*, and we can use that fact to test language comprehension. We can give the students a passage in which some of the information is missing (words or parts of words) and reasonably expect them to understand it. And if they understand it, we can also reasonably expect them to be able to replace the missing information. There are, in fact, tests which use this approach, they are generally referred to as *tests of reduced redundancy* (Spolsky, 1971; Gradman & Spolsky, 1975), and have found favour with a number of testing researchers (Cohen, 1977; Oller, 1979). This, of course, is the theoretical basis of the widely used cloze test.

3.1.1 Noise Tests

There are a number of ways of creating reduced redundancy listening tests. One of these is the noise test, in which students hear a passage which has been mutilated to some extent by the addition of background noise to the recording. This is usually "white" noise, that is noise which covers most of the frequency range of the spectrum, a sort of continuous hiss. Of course, the text is partly masked by the noise, thus reducing the redundancy of the language, but it is expected that testees will be able to use their linguistic knowledge to recreate the original text (Spolsky, 1971; Johansson, 1973; Gradman & Spolsky, 1975; Cohen, 1977). There are two main ways of recording students' responses. In the first, students are asked to record what they hear by repeating it into another microphone while they are listening to the original recording. This obviously presents practical problems for mass-testing situations such as Japanese entrance examinations. The other method is to insert pauses into the recording during which students write down what they have heard. This method would be quite practical for entrance examinations in Japan, but would obviously require students to have reasonable writing ability. We may find that some students who received low scores did so because of deficiencies in their writing ability rather than in their listening ability. In other words, the listening test scores could become contaminated by writing ability.

It seems to the writer that there are some very serious objections to noise tests. The first of these concerns their theoretical basis. The basic idea is that language is redundant and so students are expected to use their linguistic knowledge to restore the original text. However, we have established that listeners use far more than just their linguistic knowledge to comprehend, and yet the noise test seems to provide little opportunity to use nonlinguistic sources of knowledge during comprehension. Another serious objection concerns the washback effect of these tests on the classroom. Students studying for such tests will probably feel they have to spend valuable class time practising listening to recordings mutilated by random noise. This would be disastrous, as our aim is to have students learn to listen to English as it is really used for communicative purposes by English speakers. A third practical objection is that if we are going to ask students to write down what they hear, as we probably must, we are more or less giving them a dictation, in which case we might as well give them a normal dictation, which probably works just as well, is easier to make, and is far more acceptable to most teachers and students.

3.1.2 Listening Cloze Tests

Some test makers have used listening cloze tests, in which the students usually get a written passage from which certain words have been replaced by blanks, as in the standard written cloze test, and then they listen to a recording of the passage and try to fill in the blanks from what they have heard. There are a number of problems with this technique. One is that students can often treat the passage as a normal cloze test, and fill in the blanks even if they didn't hear the passage very well, in which case it is no longer a listening test at all. One way around this is to put the blanks on high-information content words, which tend to be more difficult to guess from linguistic knowledge (Henning, Gary, & Gary, 1983). But even then it doesn't make a very satisfactory listening test, as students usually find themselves simply listening for individual words rather than trying to understand and interpret a passage. In many cases such a test would be little different from a word recognition test. In terms of washback on the classroom the objections to the noise test seem to be equally applicable to the listening cloze.

There does seem, though, to be one form of listening cloze test which involves more natural listening activities and is thus likely to have a beneficial washback effect. That is to make a cloze passage on a summary of the original text. It may be argued, of course, that this is no longer a test of reduced redundancy, but that doesn't seem very important if the method produces good results. The testing procedure is quite simple. Students are given a summary of the passage they are going to hear, in which some of the important content words have been replaced by blanks. After looking at the passage for a while, so they have some idea what they are listening for, they listen to the passage. They should be discouraged from writing while listening. Their task is to use their understanding of the passage to fill in the blanks. One mark is awarded for each blank completed with an acceptable word.

Obviously, this technique would only work well with those texts which are amenable to summarizing. It seems particularly suitable for narrative texts, although it could probably be used with other text types. It avoids the problem of students just listening for single words, and if done well, could require them to use higher level processing skills, such as making inferences about the main point of the passage, the relationship between different events in the story, or the overall significance of certain parts of the passage. As the blanks are put on words which have a high information content, it is highly unlikely that students would be able to guess them without understanding the passage.

The thing that requires most care with this testing technique is making good summaries. It is not quite as easy as it seems. To those who write the summary it often seems obvious what should go into each blank, but to others this is not always the case. It is vital that summaries be checked by other teachers or experts to ensure there is no ambiguity and it is always clear what information is required in each blank (Alastair Pollitt, personal communication). Although there is little research evidence on this technique, it would seem to be one which has much to recommend it, and further research is obviously called for.

The writer has used this technique in Japan with a number of English listening tests in which the summary was written in Japanese. First, a suitable narrative text was selected. Then a number of native speakers (Japanese teachers of English) were asked to write a summary of the story in Japanese.[3] It was found that these differed a little from each other, and so a composite of these summaries was produced in order to try and get the advantages of each. Then references to those events in the narrative which were considered central to the whole story were replaced with blanks. Because of the doubt in Japanese about what exactly constitutes a word, it was decided that deletions would not be restricted to one word, instead short phrases were deleted (instructions were given on the test which made this very clear). These passages with the blanks were then given to a number of students who were asked to try to fill in the blanks without hearing the actual story, in order to check whether the blanks could be filled in from general or linguistic knowledge. It was surprising that, despite the greatest care, some blanks could indeed be guessed by a number of students without ever listening to the story, and so the summaries were remade without those blanks, and the test was finished.

The results were encouraging. Analysis showed that such tests had acceptable reliability as estimated by Cronbach's Alpha, generally comfortably over .80.[4] When one of these tests was included in a battery of comprehension tests administered to over 400 Japanese college students, it correlated closer with other tests of listening comprehension than it did with a test of reading comprehension which used the same testing method. This is an important result, because it shows that this test was functioning as a listening test, rather than just as a fill-in-the-blank test. These tests, thus, seem to have had acceptable reliability and validity.

The use of clozed summaries of listening passages as a testing method seems to have much to recommend it. They apparently have a reasonably firm theoretical base, practical trials have produced encouraging results, and they look sufficiently serious and academic for inclusion on college entrance examinations. Furthermore, the washback effect on the classroom is likely to be quite beneficial if students realize that they have to understand passages in their entirety to complete the tests.

3.1.3 Dictation

One test type which is in many ways related to the reduced redundancy tests and which has been very widely used is the dictation. This was criticized by Lado (1961, p. 34) as being little more than a test of spelling, but work in the 1970s has indicated that dictation can be a good measure of general language proficiency (Oller, 1979; Oakeshott-Taylor, 1977). Basically, the general idea is that although a passage may be redundant for first language users, it will be much less so for second language users. Thus, the extent to which second language users can utilize the redundancy in the language is a good measure of their language ability. For those who don't know the technique, dictation usually involves hearing a passage twice. The first time the passage is played straight through, and students just listen and try to understand. The second time they hear it, the passage is broken into a number of short sections with a pause between each section. During that pause students have to write down what they have just heard.

Simply listening to a passage of English and then writing it down requires the listener to engage in many of the activities we normally use in speech perception. Of course, the length of each section is very important. It requires very little language ability to repeat or write down a single word heard in isolation. However, if the length of the sections is increased, the students are required to rely less and less on their short-term memory, and more and more on their language processing ability. This is because there is a limit to the capacity of short-term memory, which seems to be restricted to about seven units of information (Miller, 1956). These units are often referred to as *chunks* and the size of these chunks is not fixed. This means that most students will be able to keep up to about seven words in short-term memory, assuming they can decode them in the first place. But more advanced students will be able to use their language ability to "chunk" these words into meaningful units, such as phrases or idea units, and then retain about seven of these units in memory. Those who have enough language ability to chunk the language can thus reproduce much longer sections of the text than those who don't. In the light of this, it is important, when constructing dictation tests, to ensure that each of the sections of text is long enough to place a certain amount of load on the short-term memory of the students. This will then force them to use their linguistic knowledge to chunk the material.

The size of each section of text between pauses can vary, but probably the best way is to make them of differing lengths, varying from about four or five words up to about 12 words or so. The ideal length will depend on the ability of the students and the nature of the text used. Experience is the only real yardstick. Ideally, the sections near the beginning of the test should be shorter, and they should get progressively longer through the test.

There are a number of ways of scoring dictations. But it is important to remember that they are not designed to be tests of spelling, and so spelling mistakes should be ignored in cases when it is obvious that the mistake is indeed a simple spelling mistake (not always an easy decision). It is common to give one mark for each correct word. This can be done by adding together all the correct words in each section. Words out of order or omitted are marked as wrong. This is an easy way to mark, but unfortunately it doesn't take account of intrusions, words written down which were not in the original passage. A better way to mark dictations, therefore, is to award students the marks for each section, one for each word in the correct version, and then subtract from that one mark for each mistake. Intrusions count as one mistake each. This could result in a minus score for some sections, in which case it is normal to award a zero.

There is a variation on the dictation test which has been used quite extensively in research on second language acquisition and in testing non-literate second language users. This is what is called a sentence elicitation task. It is basically the same as a dictation, except that students do not write down what they hear, but repeat it into a tape recorder during the pause after each section. This has shown itself to be a reasonably reliable method of testing general language ability (Gallimore & Tharp, 1981), but it suffers from the disadvantage of not being very suitable for the sort of large-scale testing which is necessary in Japan. It also seems to be unlikely to lead to a good washback effect on the classroom, as students will likely spend long periods of time just repeating sentences after a tape recording.[5]

Dictations are now firmly established in the second language teaching world as good tests of general second language ability, which are reasonably easy to make, not too difficult to mark and are generally reliable and reasonably valid.[6] Dictations obviously require students to write, and listening scores may be influenced by second language writing ability in cases where students have widely different levels of written English. However, the writer has administered a number of English dictations to college students in Japan, and has found that most students have sufficient writing ability to handle dictations quite well. Dictations also seemed to cause far less student animosity than cloze tests. Furthermore, they seem to be considered quite acceptable academic tasks by the majority of teachers. However, if we want students to spend their classroom time listening to realistic communicative spoken language, then their washback effect will probably not be as positive as we would like. Students will likely practise listening to decontextualized written texts read out aloud, which is very different from listening to communicative spoken language. With this one very serious reservation, dictation does seem to be a possible candidate for inclusion in entrance examinations.

3.2 The Product Approach

As we can now see, the main problem with testing the *process* of listening comprehension is that we don't really know enough about it. An obvious alternative approach is to test the *product* of listening. We can give students a passage to listen to and then see if they understood it. The drawback with this approach is that the product of the listening comprehension process is not easily available for inspection. It lies inside the student and we cannot record it, or take it home for grading. Assessing listening ability is an indirect process. We have to give students some task which we think is dependent on comprehending a piece of language, and then try to infer from that whether we think they understood the language or not. The task most commonly given to students is answering questions, but it could also be filling in a grid, or marking a place on a map or chart. Picture identification tasks are also sometimes used.

Naturally, it is important that completing the task is dependent on understanding the text. Care should be taken that task completion really does require comprehension of the text. It is surprising how often a good guess can be made at the right answer without even hearing the actual text (Preston, 1964; Connor & Read, 1978).

3.2.1 Intervening Variables

There is, then, no such thing as a "pure" test of listening. We have to use indirect measures of listening, and when we choose a test format, it is important to realize that answering the questions will require other abilities apart from those we want to measure. This, of course, will confound the measurement of listening comprehension with other variables which we may not be interested in measuring at all. For instance, if the questions on an English second language listening test are presented to the student in written English, then ability to answer the questions will depend not only on second language listening ability, but also on second language reading ability.[7] A student with very slow reading may fail to answer many of the questions because he couldn't read fast enough, even though he may have understood the spoken text quite well.

One possible way around this problem is to give the question in oral form at the same time as the text is presented. The students listen to a passage, recorded on a tape, and then listen to the questions, which are also recorded on the tape.[8] This solves the problem of L2 reading ability contaminating listening tests, but introduces a different problem, namely that some students have better memories than others. Students with poor memories will obviously suffer. This might not be such a problem if the questions are rather short, and not such a strain on memory, but it could be a serious problem with multiple choice questions. Such tests also call for a high level of concentration from the test taker, as one short moment of inattention can result in the student not hearing the question and being unable to answer despite the fact that he understood the text quite well.

To some extent these problems can be avoided by asked open-ended questions, which would require less reading or less memorization. However, if the answers are required in the second language, then, of course, the listening comprehension test score will probably be contaminated by writing ability. One simple solution, at least in cases where all the students share the same first language, as in Japan, is to ask questions using the students' first language, and also allow them to answer in their first language. Although students obviously vary in their first language abilities, the difficulty of the Japanese used can be kept to such a basic level that it is unlikely to tax the ability of even the least proficient student. This avoids many of the problems mentioned above, but moving backwards and forwards between two languages can cause problems, and some students find it irritating.

The choice of question form will depend largely on the purpose of the test, and the priorities of the test maker. Some teachers, for example, may be quite happy to have questions written in English, and thus have their listening tests contaminated by student reading abilities. After all, listening and reading are both important language skills, which are obviously closely related, and a test which measured both could well be a good test of general language comprehension. In the case of test instructions, though, the problem is rather more serious, as the penalty for misunderstanding one simple instruction can be the loss of many marks on the test. A penalty usually far out of proportion to the mistake made, which would suggest that, wherever possible, instructions should be in the students' first language.

The writer has used both multiple choice and open-ended questions, written in Japanese, on English listening tests and has found they work quite well if a little care is taken. On the open-ended questions, students were given the option of answering in either Japanese or English, which seemed the best solution, and results were generally satisfactory. Furthermore, using Japanese questions enables the test writer to ask questions which are more complex, and hence probe deeper into the students' understanding of the passage.

3.2.2 A Purpose for Listening

In "real life," when we listen to anything, we virtually always listen with some purpose, even if the purpose is something as vague as general interest. If we want to replicate real-life listening, it is very important to give students a purpose for listening. This also relieves them from the necessity of trying to remember all the information in the passage, which is a very unnatural thing to do (and generally impossible), although it is quite common for students to try to do so in important testing situations. There are a number of ways the listening purpose can be set up, but perhaps the simplest is just to give students the questions before they hear the passage. A somewhat more interesting alternative is to give specific instructions about the listening situation and the purpose of their listening, a sort of listening role-play, and then allow students to select the relevant information for themselves.

3.2.3 Authenticity

Another important consideration is the nature of the texts used. There is currently a very strong trend in language teaching to provide students with "authentic" language. While it is often difficult to be specific about exactly what we mean by authentic (Besse, 1981), the trend is an admirable one. Obviously, if we want our students to learn a language, we must give them realistic language to learn, and have them engage in realistic language activities with it. Of course, in one sense, the language learning classroom itself is not really an authentic situation, and the testing situation is even less an authentic communicative situation than the classroom. Whatever we try to get our students to do, however much we try to set up authentic situations, our students' prime concern will be to get high marks on the test. Nevertheless, given the tremendous washback effect of the test on the classroom, test makers owe it to teachers and students to use texts which are as authentic as possible, in order to encourage students to study authentic texts in the classroom.

3.2.4 Making Multiple Choice Tests

In most standardized tests, such as the TOEFL or TOEIC, it is normal to use multiple choice questions to test comprehension. These are very convenient to score with modern high-speed marking machines, but making them, however, is not so easy. Not only do we have to write a sensible question and provide the correct answer, but we must think of three distractors, or alternative answers, which must look like attractive possibilities to those who did not understand the text, but which are clearly wrong to those who did. This is extremely difficult to do. Multiple choice items are very complex, and it is often difficult to tell how they are going to work by just looking at them. Therefore it is necessary to pre-test them on students similar to those who are to take the test, and then subject the results to statistical analysis in order to find out which items are good and which are not. Bad questions, or distractors which are not distracting any of the students, can then be rewritten; and it may often be necessary to reject or rewrite large numbers of the items. Using such a method, it is quite possible to produce tests which have very satisfactory statistical properties. However, pre-testing important entrance examination questions in

Japan is generally considered impossible, which suggests that test makers should avoid using multiple choice questions.

3.2.5 Open-ended Questions

Of course it is possible to have comprehension questions without using a multiple choice format. One of the simplest ways is just to ask open-ended questions. Students then write the answer in their own words. The most obvious problem is that we then have to mark the answers, and that clearly involves deciding what is to be marked right and what wrong. Such judgments can be very difficult when the student has produced something which seems almost right, but which we feel is in some way inadequate. A related problem is how to deal with English answers that are grammatically incorrect. It seems reasonable to ignore small mistakes, but if the mistakes are such that they cause ambiguity, we have the problem of deciding what the student meant to say. It is necessary to get the criteria very clear, and ensure that all markers are applying the same standards. This problem can be avoided to some extent by using short-answer questions, in which students have to give very short answers of only one or two words. Such question types are ideal in the case of texts such as public announcements, or other information-style texts, where the main point of the message is clear and thus what answers are acceptable and what are not. Furthermore, the washback is likely to be very good, as such passages are a type of text we would like our students to study in the classroom. They are, after all, just the sort of thing they will have to understand if they go to an English-speaking country. Such short-answer questions are quite easy to make, students only spend a short time writing the answers, they are easy to mark, and answers can often be given in Japanese. In fact, they would seem to be ideal candidates for inclusion on college entrance examinations.

There are a lot of texts, though, which are not suitable for such short-answer questions. Sometimes teachers want to test more than just the ability to extract clearly stated information; they may also want to examine whether students have understood the main point of a passage, can produce a summary, or can make inferences about the events in the text. In such cases longer answers can produce far more information. However, the problem of deciding which answers are suitable and which are not becomes far more acute, as responses will vary on a continuum from completely satisfactory to totally inadequate. The markers will often have to use a scale to rate the suitability of responses, and this is a far cry from simply marking items right or wrong.

Some sort of rating scale has to be developed, which need not be too complex, but which will certainly take some time and effort to produce. Exactly how this is made will depend on a number of factors, but basically it is necessary to decide what answers students are likely to produce, decide how these should be evaluated relative to each other, and decide how many marks to award for each. Having developed a scale to rate the answers, the job has only begun. It is then necessary to find out if the scale works on sample test items, after which it will probably be necessary to modify the scale a litte. Next, raters must practice using the scale until they can all apply it consistently. This last step is very important. Students whose papers were marked by a generous rater have a far higher chance of passing the examination than those whose papers are marked by a stricter rater. Even the same rater tends to vary in the way he applies a rating scale over a period of time. The rater may get more severe, or less severe, or start paying more attention to one aspect of the scale than before. This variation in application of the rating scale is not only unfair, but can turn the examination into a worthless lottery.

3.2.6 Grids, Diagrams and Pictures

Because of the problems associated with using questions in comprehension tests, some test makers have suggested trying to avoid these all together, by using diagrams or pictures (Heaton, 1975; Brown & Yule, 1983). Here again, though, there are a number of things to be considered and options to be weighed. First, while it is easy to use these to test simple or short samples of language, it is often not so easy to use them to test more complex language or extended discourse. Diagrams, and even more so, pictures, can take a considerable amount of preparation time, and often require drawing skills which many of us don't possess. Test makers who are good at drawing may enjoy making picture recognition tasks, in which students have to choose a picture to match a description they hear, but most of us are not very good at drawing and have to search around in old textbooks and such places to try to find pictures which we can use. Although picture recognition tasks may not look very serious or academic, they are used on the

TOEIC, which is a generally respected examination.

However, diagrams do offer possibilities of testing realistic language in a number of important situations. One obvious possibility is to give students a map and have them follow a conversation in which directions are given. Students follow the directions and then indicate on their answer sheet what the destination was. Another possibility is to give students an incomplete diagram, or a picture, and have them complete it by listening to a description of it. This sort of activity is very good for the language classroom, and can be enjoyable and extremely useful, but again, somehow, one wonders if they would be considered appropriate for entrance examinations. Perhaps a compromise could be found in filling in grids, which are something like a halfway point between short-answer questions and diagrams. Grids can be used for a host of activities, almost anything which involves timetables, or classifications (of objects or countries or blood types or whatever). They can be made such that almost no reading ability or memory is involved, and very little writing.

Almost all the activities mentioned in this section could lead to quite a good washback effect on the classroom, encouraging students to listen to railway announcements, weather forecasts and all sorts of other realistic and useful things. It would be very nice to see some of these tasks used in entrance examinations.

4. Conclusions

The first conclusion to note is that testing listening comprehension is far more complex an undertaking than it would seem to be at first sight. However, this applies not only to testing listening comprehension, but to testing any aspect of second language ability. Language use is a very complex phenomenon, about which far too little is known. The question test makers have to ask themselves is whether they are prepared to face up to the complexity of the issue or not. The writer believes very strongly that, as professionals, there is only one answer we can possibly give to such a question.

This article has discussed eleven test types. Table One lists the types, comments upon them, and estimates their likely washback effects.

In practical terms, the choice of test type will depend to some extent on the needs of the situation and the opinions of individual teachers. Of the various test types listed in Table One, cloze passages on summaries of texts seem very suitable for entrance examinations and are obviously highly recommended, especially with narrative texts. Short-answer questions on realistic public announcements are another obvious choice and should be seriously considered. For expository tests, more complex open-ended questions can be used, or grid-filling activities where suitable. Grids or picture selection tasks are very suitable for descriptive texts. There is no reason why questions and answers should not be in Japanese. The writer feels every entrance examination could benefit from a question in which students hear directions on how to get to a destination and then find the destination on a map. Dictation is a possibility, or rather a temptation, because, although it is obviously a good testing technique, the likely washback effect argues against it.

The most important thing to bear in mind when making college entrance examinations in the Japanese situation is the washback effect on the classroom. If test makers continually ask themselves what sort of washback their test will lead to, there is every possibility of improving English language education in Japan.

Table One
Summary of Listening Comprehension Tests

Test	Comments	Washback
Noise tests	Utilizes little non-linguistic knowledge. Tape-recorded responses are impractical. Dictation is probably just as good, and more acceptable.	Very negative.
Ordinary listening cloze	Can often be done without listening. Students often just listen for individual words.	Probably not very positive.
Cloze on *summary* of passage	Students must understand whole passage. Can test inferencing and higher order skills. Summaries must be carefully produced, but summaries in Japanese are possible. Highly recommended.	Likely to be very positive indeed.
Dictation	Requires L2 writing ability. Well established testing technique. Easy to make and administer.	Rather negative.
Sentence elicitation	Same as dictation except that students record responses into a tape recorder. Rather impractical for large numbers. Good for non-literate students.	Rather negative.
Multiple choice comprehension questions	Can be machine-scored. Difficult to make, require pre-testing and item analysis.	Depends on questions made.
Short answer comprehension questions	Easy to make, and not too difficult to mark. Can be used with realistic communicative texts. Suitable for questions and answers in Japanese. Highly recommended.	Very positive if used on realistic texts.
Open-ended longer questions	Easy to make, but usually require a rating scale to mark them. This takes time and coordination between raters. Can be used to test understanding in depth.	Quite positive if used on realistic texts.
Picture recognition tasks	Requires a lot of time or drawing skill to make. Take up a lot of space on test paper. Don't look very academic.	Good if natural English and realistic tasks are used.
Diagram completion tasks	Easier to make than pictures. Can be used to test natural language use. Following street directions is a very good communicative activity easily incorporated in any test. Don't look very academic.	Good if natural English and realistic tasks are used.
Grid completion	Easy to make. Can be used to test natural language. Endless possibilities, especially with timetables and classifications.	Good if natural English and realistic tasks are used.

Notes

1. It is a basic truism of educational measurement that if a test is too difficult for the students on which it is used, it will actually give very little information regarding their ability.
2. The writer does not feel a need to distinguish between the terms acquisition and learning. They will be used interchangeably in this paper.
3. I would like to take this opportunity to offer thanks to the teachers at Osaka Meijo Women's College who kindly cooperated in this study.
4. Cronbach Alpha makes the assumption that the items in the test are independent of each other. If they are not, the reliability estimate appears higher than it should be. It is possible that the items in this test, as with all cloze tests, don't meet this requirement in all respects and thus the reliability could be a little lower.
5. This might be acceptable to those who feel that the Audio Lingual Method is a good way to learn languages, but of course, such a view is now very rare among applied linguists.
6. It should be noted that there is no such thing as a testing method which automatically produces reliable and valid tests. Unfortunately, it is necessary to validate each new test, and even each new use of an old test.
7. A look at TOEFL practice listening tests will show just how much reading is actually involved in some listening tests.
8. Both the TOEIC listening section, and the JACET-COLTD Listening Comprehension Test have questions recorded on the tape.

References

Alderson, J. C., & Urquhart, A. H. (eds.). (1984). *Reading in a foreign language.* London: Longman.

Besse, H. (1981). The pedagogic authenticity of a text. In *ELT Documents special: The teaching of listening comprehension.* London: The British Council.

Brown, G., & Yule, G. (1983). *Teaching the spoken language.* Cambridge: Cambridge University Press.

Bruce D. J. (1985). The effects of listeners' anticipations on the intelligibility of heard speech. *Language and Speech I*, 79-97.

Buck, G. (in press). *Written pronunciation tests: Do they work?* ELT Journal.

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

Cohen, A. (1977). Redundancy as a tool in measuring listening comprehension. In R. Dirven (ed.), *Horverstandnis im Fremdsprachenunterricht. Listening comprehension in foreign language teaching.* Kronberg/Ts.: Scriptor.
tests. *Language Learning 28* (1), 149-157.

Devine, J., Carrell, P. L., & Eskey, D. E. (eds.). (1987). *Research in reading in a foreign language.* Washington, D.C.: TESOL.

Dirven, R., & Oakeshott-Taylor, J. (1984). Listening comprehension (Part I). *Language Teaching 17*, 326-342.

Dirven, R., & Oakeshott-Taylor, J. (1985). Listening comprehension (Part II), *Language Teaching 18*, 2-20.

Faerch, C., & Kasper, G. (1986). The role of comprehension in second language learning. *Applied Linguistics 7*, 257-274.

Gallimore, R., & Tharp, R. G. (1981). The interpretation of elicited sentence imitation in a standardized context. *Language Learning 31* (2), 370-392.

Goodman, K. (1970). Reading: A psycholinguistic guessing game. In H. Singer & R. B. Ruddell (eds.), *Theoretical models and processes of reading.* Newark, Del.: International Reading Association.

Gradman, H. L., & Spolsky, B. (1975). Reduced redundancy testing: A progress report. In R. L. Jones & B. Spolsky (eds.), *Testing language proficiency.* Arlington, Va.: Center for Applied Linguistics.

Grosjean, F. (1980). Spoken word recognition and the gating paradigm. *Perception and Psychophysics 28*, 267-283.

Heaton, J. B. (1975). *Writing English language tests.* London: Longman.

Henning, G., Gary, N., & Gary, J. (1983). Listening recall: A listening comprehension test for low proficiency learners. *System 11*, 287-293.

Johansson, S. (1973). An evaluation of the noise test: A method for testing overall second language proficiency by perception under masking noise. *IRAL 11*, 107-131.

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

Lado, R. (1961). *Language testing.* London: Longman.

Leech, G. N. (1983). *Principles of pragmatics.* London: Longman.

Marslen-Wilson, W. D. (1980). Speech understanding as a psychological process. In J. C. Simon (ed.), *Spoken language generation and understanding.* Dordrecht: Reibel.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review 63* (2), 81-97.

Oakeshott-Taylor, A. (1977). Dictation as a test of listening comprehension. In R. Dirven (ed.), *Horverstandnis im Fremdsprachenunterricht. Listening comprehension in foreign language teaching.* Kronberg/Ts.: Scriptor.

Oller, W. J. (1979). *Language tests at school.* London: Longman.

Preston, R. C. (1964). Ability of students to identify correct responses before reading. *The Journal of Educational Research 58* (4), 181-183.

Richards, J. C., & Rodgers, T. S. (1986). *Approaches and methods in language teaching.* Cambridge: Cambridge University Press.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (eds.), *Theoretical issues in reading comprehension.* Hillsdale, N.J.: Lawrence Erlbaum.

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (eds.), *Schooling and the acquisition of knowledge.* Hillsdale, N.J.: Lawrence Erlbaum.

Savignon, S. J. (1983). *Communicative competence: Theory and practice.* Reading, Mass.: Addison-Wesley.

Schank, R. C., & Abelson, R. P. (1977). Scripts, plans, and knowledge. In P. N. Johnson-Laird & P. C. Wason (eds.), *Thinking: Readings in cognitive science.* Cambridge: Cambridge University Press.

Smith, F. (1985). *Reading.* Cambridge: Cambridge University Press.

Spolsky, B. (1971). Reduced redundancy as a language testing tool. In G. E. Perren & J. L. M. Trim (eds.), *Applications of linguistics.* Cambridge: Cambridge University Press.

Spolsky, B. (1975). Language testing: Art or science? Paper presented Aug. 27, 1975 at the Fourth AILA World Congress, Stuttgart, Germany.

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: $13/yr.
(must prepay)
Institutions: $30/yr.
Foreign subscribers add $3/yr. for regular postage, or $12.50/yr. for Airmail delivery.

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

Cambridge
ELT

New books for teachers

Understanding Research in Second Language Learning

A teacher's guide to statistics and research design

James Dean Brown

Designed specifically for language teachers with no previous background in statistics, this book focuses on the skills and processes necessary for understanding and critically reading statistical research in language learning. Brown explains the basic terms of statistics; the structure and organisation of statistical research reports; the system of statistical logic; and how to decipher tables, charts and graphs.

'This book does an excellent job of filling a major gap in the publications available in our field.'
Lyle F. Bachman, University of Illinois at Urbana-Champaign

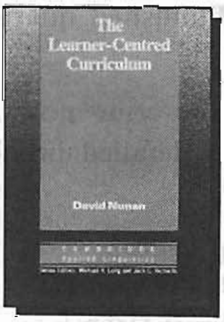
New Directions in Language Teaching

The Learner-Centred Curriculum

David Nunan

Traditionally, the curriculum has been regarded as the statement of what *should* be done in a course of study. This book takes as its starting point what *is* done by language teachers in their classes. The concept of a negotiated model is developed, in which the curriculum is arrived at collaboratively between teachers and learners.

The author looks at the curriculum from the teacher's perspective, and reports on what teachers focus on in the planning, implementation and evaluation of language courses.



Cambridge Applied Linguistics

For further information on all Cambridge ELT publications, please contact
Steven Maginn, Cambridge ELT Office,
c/o U.P.S. Ltd., Kenkyu-sha Bldg.,
9 Kanda Surugadai 2-chome, Chiyoda-ku,
Tokyo 101. Tel: (03) 295-5875.

CAMBRIDGE UNIVERSITY PRESS

JALT Journal, Vol. 10, Nos. 1 & 2

YAKUDOKU: JAPAN'S DOMINANT TRADITION IN FOREIGN LANGUAGE LEARNING

Nobuyuki Hino

Abstract

The mainstream of the teaching of English in Japan is *yakudoku.* In this method, English is first translated into Japanese word-by-word, and then the resulting translation is reordered to match Japanese word order. In Japan, the learning of the yakudoku technique is often identified with the goal of studying English itself. In fact, yakudoku is a deeply rooted sociolinguistic tradition in Japan, which dates back over a thousand years to when the Japanese started to study Chinese. This paper examines the nature of yakudoku by investigating how it has continued to be used in foreign langauge learning in Japan. The implication of the yakudoku tradition for the teaching of English today is also discussed.

1. Introduction

In choosing a teaching method for an EFL country, it is important to investigate the methodological antecedents. That is, those who try to improve the English language education of an EFL country should not ignore its indigenous educational traditions. Otherwise those traditions may become obstacles to the diffusion and implementation of the innovations which the reformers promote (Henrichsen, 1988).

As was briefly described by Hino (1982), Japan has a long tradition of foreign language learning called *yakudoku* (訳読). In the present paper, the sociolinguistic nature of yakudoku will be examined with an emphasis on its historical aspects, and its implication for the teaching of English today will be discussed.

Nobuyuki Hino is Associate Professor of EFL at Tokyo International University. He has published in *JALT Journal, World Englishes, Georgetown University Round Table on Languages and Linguistics,* and has presented at TESOL and JALT conventions.

2. What is Yakudoku?

Yaku (訳) means "translation," and *doku* (読) means "reading." Yakudoku is defined as a technique or a mental process for reading a foreign language in which the target language sentence is first translated word-by-word, and the resulting translation reordered to match Japanese word order as part of the process of reading comprehension (Kawasumi, 1975). An illustration of the yakudoku process might be as follows:

[Target language sentence] She has a nice table in her room.

Stage I [The reader mentally makes word-by-word translation.]

She	has	a	nice	table	in	her	room
kanojo	motteiru	hitotsu-no	sutekina	teburu	naka	kanojo-no	heya

Stage II [Translation reordered to match Japanese syntax.]

kanojo kanojo-no heya naka hitotsu-no sutekina teburu motteiru

Stage III [Recoding in Japanese syntax.]

Kanojo-wa kanojo-no heya-no naka-ni hitotsu-no sutekina teburu-wo motteiru.

There are two aspects to yakudoku. One is the regressive eye movement resulting from the word-by-word translation. The other is the fact that the meaning is not understood directly in the target language but only via translation (Ueda, 1979).

In teaching students how to read English, teachers introduce students to the yakudoku technique. This may be referred to as the Yakudoku Method of teaching English, the method used by the majority of Japanese teachers of English from junior high to college level. The Yakudoku Method aims to teach yakudoku skills to a stage where the student is able to use the technique without help from the teacher. The teacher's job in class is to explain the word-by-word translation technique, to provide a model translation, and to correct the student's translation (Kakita, 1978; Tajima, 1978).

Two nation-wide surveys conducted by the Japan Association of College English Teachers (Koike et al., 1983, 1985) showed that from 70 to 80 per cent of Japanese teachers of English in high schools and universities resorted to the Yakudoku Method. Hino (1987a) also found that 70 per cent of his university students had been taught to read English solely with this method. Yakudoku is "the" method in the teaching of English in Japan.

3. Disadvantages of the Yakudoku Method

For many Japanese students, reading English and yakudoku are the same thing (Matsumoto, 1965; Tazaki, 1978). They are neither aware that it is much more natural to read English in the original word order nor that it is desirable to read directly in English without recourse to a Japanese translation. Hino (1987a) observed that Japanese students of English tend to use the word *yakusu* (= translate) synonymously with *yomu* (= read). Having been trained to read English via translation, they have come to identify this with the process of reading in a foreign language itself. It may even be that the goal of reading a foreign language text is regarded as simply to render it into a possible Japanese equivalent, without consideration of the value of the translation in understanding the contents of the original. Once the English is transformed into Japanese, it is considered read (Tazaki, 1978; Osawa et al., 1978). Conversely, if an English text has not been recoded into Japanese, "reading" is not considered to have taken place (Ueda, 1979; Kakita, 1978). The yakudoku habit clearly is a severe handicap for the Japanese student. It limits the speed at which the student reads, induces fatigue, and reduces the efficiency with which s/he is able to comprehend. The meaning of a text is obtained via Japanese translation, and is only an approximation to the original.

Yakudoku also has detrimental effects on the other language skills — listening, speaking, writing. Students who have been trained in yakudoku reading employ a similar strategy in listening comprehension. They attempt to understand speech by translating every sentence into Japanese (Tazaki, 1978). As a consequence, they cannot follow speech unless it is delivered slowly, and they find comprehension a tiring, imprecise, and ineffective process. In speaking and writing, the yakudoku process is applied in reverse. A Japanese sentence is composed, translated into English word-by-word, and then the words are reordered according to English syntax (Matsumoto, 1965). The result is seldom idiomatic English sentences, and is produced very slowly.

In spite of these serious disadvantages, why is the Yakudoku Method so prevalent? It is important to note that the Course of Study for English prescribed by the Ministry of Education which defines and controls the contents of English teaching in junior and senior high schools, makes no mention of the skill of translating English into

Japanese (cf. Mombusho, 1978a, 1979a). In other words, the Education Ministry by no means encourages yakudoku. The Yakudoku Method of teaching English is not necessarily something that is politically imposed upon the teachers by the administration, but is a long established tradition which exists at a deeper level of the sociolinguistic structure of Japan.

4. History of Yakudoku

Yakudoku as a method of reading has a long history, though exactly when it started is still unknown. It goes back more than a thousand years (Suzuki, 1975), when the Japanese began to read Chinese, that is, the first foreign language they studied. They read Chinese by translating it into Japanese word-by-word. This process is basically the same as the current yakudoku practice. For example:

[Target language sentence] 毎見秋瓜憶故丘[1]

Stage I [Word-by-word translation]

見 秋瓜 憶 故丘

goto miru shuka omou kokyu

Stage II [Reordering]

shuka miru goto kokyu omou

Stage III [Recoding in Japanese syntax]

Shuka-wo miru goto-ni kokyu-wo omou

However, there is an important difference between the original form of yakudoku and the current yakudoku practice in reading English. Today, yakudoku is usually an implicit mental process. Yakudoku in reading Chinese, on the other hand, was an explicit process. At Stage II, some symbols are added to indicate the Japanese word order:

毎レ見二秋瓜一憶二故丘一

The symbol , for example, indicates the reversal of the two adjoining characters. Symbols and above came from Chinese numerals, but are used here as signs which direct the reordering according to a set of rules. At Stage III, Japanese postpositions and suffixes are written in *katakana* beside the Chinese words:

毎ニレ見ル二秋瓜ヲ一憶フ二故丘ヲ一

Even today, this method of reading classical Chinese is taught in senior high schools in Japan as part of the instruction in the Japanese language, following the Course of Study for the National Language issued by the Ministry of Education (cf. Mombusho, 1978b, 1979b).

Having perfected the yakudoku technique, in later years the Japanese came to apply it to the study of other foreign languages. In the 19th century, the Japanese produced textbooks for the study of Dutch in which the yakudoku technique was used. *Kunten Oranda bunten* (1857) is a typical example of the application of yakudoku to Dutch, with the word *kunten* in the title meaning the symbols used for the reordering. For each Dutch sentence in this text, the Japanese equivalents of the Dutch words are written out in *kanji* with postpositions and suffixes added in katakana, which are to be reordered according to the same symbols as the ones used in reading Chinese. For example:

諸ノ本名ガ　之　婦人
Alle-eigennamen van vrouwen
二　上　レ

(Reprinted in Sogo, 1970)

After Dutch, yakudoku was applied to the study of English. The following is an excerpt from *Eibei taiwa shokei* (A shortcut to English conversation, 1859), an English textbook written by Manjiro Nakahama (1827-1898) toward the end of the Edo or Shogun Period. The Japanese words in *hiragana* are word-by-word translations, which are to be reordered into the Japanese word order in accordance with the reordering symbols:

わたくし　よろこぶ　ことを　みる　おまんの　おけるを　よき　うまきことに
レ　レ　三　二　一
アイ　アム　ハペ　ツ　シー　ユー　イン　グーリ　ヘルス[2]
I am happy to see you in good health.

(Reprinted in Kawasumi, 1975)

In the Meiji Period, the reordering symbols were replaced by numerals, a simpler way of indicating the Japanese word order. Still, the process of word-by-word translation and reordering itself remained the same. Below is an excerpt from *Soyaku Rigaku Shoho* (First lessons on natural philosophy with Japanese translations, 1871), an English text which in the modern term would fall into the category of content-based language instruction. The Japanese equivalents for each English word in kanji and katakana are numbered by kanji numerals.

What hemisphere is America in?
ドノ　半球ニ　有カ　アメリカガ　於テ
一　二　五　四　三

(Reprinted in Sogo, 1970)

The explicit writing of numerals also gradually came to be less frequently used in the reading and teaching of foreign languages. However, the essential process of yakudoku, though more implicitly, continued to be widely practiced. Yakudoku is a strong educational or sociolinguistic tradition in Japan, which enjoys a history of over a thousand years. But that is not to say that it has been without its critics.

5. Criticisms of Yakudoku

An early critic of yakudoku was the Confucianist Sorai Ogyu (1666-1728) with his disciple Shundai Dazai (1680-1747) (Suzuki, 1975). Ogyu voiced objections to yakudoku in his book *Gakusoku* (Rules of learning) written in 1727. His main point may be summarized as follows:

> The traditional method of reading Chinese is a misleading one, which should be avoided. You cannot truly understand Chinese in this way. Chinese should be read as Chinese.
> (Kawasumi, 1975)

Ogyu warned that the spirit of the Chinese people could not be grasped through yakudoku, which is merely a literal translation that ignores the linguistic and cultural differences between the two languages.

Gentaku Otsuki (1757-1827), scholar of the Dutch language and culture, criticized the application of yakudoku to Dutch in his book *Rangaku Kaitei* (Steps in Dutch studies) written in 1788. His main argument is similar to Ogyu's criticism:

> In reading Dutch, beginners may use the method used for reading Chinese, but it is desirable for advanced readers to read directly in the original word order. You can understand the contents more clearly this way. Dutch often loses its meaning if rendered into Japanese.
> (Kawasumi, 1975, 1978)

It is recorded that some people, though very few, were able to read directly in Dutch. The 19th century scholar Genichiro Fukuchi was one of them. He describes his experience and that of two other non-yakudoku readers:

> In those days, most people read Dutch in the same way as they did Chinese. Mr. Seikyo Sugita read, however, directly in Dutch without translation. Others achieved understanding only via word-by-word translation. I also read directly in Dutch, though I may sound arrogant.... I insisted that Dutch should not be read via such forward

> and backward translation.... Rinsho was just 14 years old, but he read Dutch in my way. He was able to read three times as fast as the other students.
> (Reprinted in Kawasumi, 1978. Translation the present writer's.)

Though Fukuchi said "in those days," the situation actually remains unchanged even today. A large number of Japanese believe yakudoku to be the normal way to read a foreign language.

In 1911, Yoshizaburo Okakura of Tokyo Koto Shihan Gakko (presently the University of Tsukuba) published a book entitled *Eigo Kyoiku* (English language education). This was the first systematic study of the teaching of English in Japan. Here, we find a thorough criticism of yakudoku:

> In the teaching of English in our country, students are taught to translate word-by-word, with forward and regressive eye movement. This is a strongly established convention. I think this comes from our traditional method of reading Chinese, in which Chinese words are reordered to match Japanese word order.... This is a wrong method, which treats Chinese not as a foreign language, but as a kind of Japanese. We should not use this method in studying English.... It is a pity that everyone considers this to be the only way of reading foreign languages.
>
> In reading Chinese, it is best if you understand the meaning of a text in the original word order. The contents are understood well enough in this way. As a matter of fact, this is the best way to achieve understanding. Likewise, direct reading is the best way of reading English in terms of time, energy, and efficiency.
> (Reprinted in Kawasumi, 1978. Translation the present writer's.)

Today, criticisms of yakudoku are frequently found in Japanese TEFL journals as well as in books and magazines for the general public. In spite of these criticisms, yakudoku still dominates the way Japanese read foreign languages. It dies hard.

6. Why is Yakudoku so Persistent?

Why is yakudoku persistently practiced? As a cause of the widespread practice of yakudoku, many analysts refer to its easiness for the teacher (Tazaki, 1978; Ozeki et al., 1983; Ito, 1984). That is, the use of the Yakudoku Method requires little professional training, and also little preparation is needed for each class. Anyone who has studied English through yakudoku is able to teach it in the same way without much effort. The inadequate training system of EFL teachers in Japan enhances this tendency (Ozeki et al., 1983; Hino, 1987a). Not having been exposed to alternative approaches, many teachers are liable to depend on the same old method with which they have been taught.

Mental discipline is also often cited as a major function of yakudoku (cf. Hiraizumi & Watanabe, 1975). The decoding and deciphering activities involved in the yakudoku process provide the learner with opportunities for mental training, and thus make the Yakudoku Method worthwhile.[3]

Though these analyses explain some aspects of the persistence of yakudoku, they are not necessarily considered to be crucial from the standpoint of the present paper. The fundamental nature of the yakudoku phenomenon is more sociocultural than pedagogical.

Once a practice is accepted as a tradition, it becomes a norm. No one is accused as long as s/he follows this norm. On the other hand, those who do not observe the norm are regarded as deviants. And the longer history the tradition has, the stronger the norm is. The society gets the tradition rolling in accordance with the law of inertia. This tendency is especially conspicuous in a rigidly structured society such as Japan. This is the sociocultural view of the survival of yakudoku that the present article is based upon.

Yakudoku was an effective method when Japanese first faced classical Chinese. The unfortunate fact is that yakudoku has become such a strong convention that Japanese find it very difficult to free themselves from this tradition even in the 1980s when English needs to be learnt as a means of communication rather than as a dead language.

7. Conclusions

In developing or selecting teaching methodology suitable for an EFL country, it is essential to investigate its indigenous sociolinguistic tradition. In the case of Japan, it is the yakudoku tradition of learning foreign languages, which has survived over a thousand years and is still alive and well. Yakudoku may even be viewed as the solid infrastructure of language teaching in Japan. In the past, yakudoku may have had its own value as one possible method of studying foreign cultures (Ito, 1978, 1984; Ozeki et al., 1983). However, in terms of the teaching of English for communication needed today, it is undoubtedly a serious handicap for Japanese students of English.

For years, many foreign teachers have been trying to introduce new, innovative methods to the teaching of English in Japan, but few of them have enjoyed remarkable success. It is easy to imagine that those teachers from abroad were not adequately informed of the yakudoku tradition. In fact, it may be argued that no new method could succeed in Japan unless it penetrates into this deeply rooted convention which governs the Japanese student as well as the Japanese teacher.

Pedagogy for teaching English in Japan should put its first priority on helping students overcome the yakudoku habit. In fact, a growing number of EFL teachers are making such efforts in Japan. Kasajima (1987) presents several techniques to achieve direct understanding in English, including sense-group reading which discourages the regressive eye movement caused by word-by-word translation. Sagawa and Furuya (1984) make extensive use of timed reading, using materials with limited vocabulary and comprehensible contents. Hino (1987b) suggests that comprehension practice in listening to English spoken at a normal speed would be effective in eliminating the yakudoku habit. Based on their own experience in learning English, some successful Japanese learners of English claim repeated reading aloud to be a good way to achieve direct understanding in English (Kunihiro, 1970, among others). Software for personal computers which integrates timed reading and word-group reading has also been developed. However, the use of these various approaches is still restricted to the minority of Japanese teachers of English.

Lastly, it should be explained that this paper is by no means another attempt to promote the misleading "Japan is unique" theory which is believed by some Japanese. In fact, the practice of the Yakudoku Method is not necessarily specific to Japan. The Grammar-Translation Method in the West, which grew out of the teaching of classical languages such as Latin and Greek, presents a close resemblance to the Yakudoku Method. In the East, Deyama (1988) reports that translation exercises similar to yakudoku play the central part in the teaching of English in Korea. Kanno (1986) even suggests that Koreans in their early history may also have used reordering symbols of their own when they encountered the Chinese language.

Yakudoku is not exactly a problem peculiar to Japan. Still, the fact remains that this tradition, which enjoys a long history, is a hard habit to break for the Japanese. It is necessary for all EFL teachers in Japan to be well acquainted with the nature of yakudoku.

[An earlier version of this paper was presented at the 1987 JALT Conference, Waseda University, Tokyo. I am indebted to many scholars for their valuable comments and suggestions, including Professors Jack Richards, Ted Plaister, Kenneth Jackson, Lynn Henrichsen, Barry Duell, Eric Morrison, and Madoka Ogiwara.]

Notes

1. 毎= every time, 見= see, 秋瓜= autumn melon, 憶= remember, 故丘= hill at my home; "Every time I see an autumn melon, I remember the hill at my home."
2. These *katakana* letters are the transcriptions of the English pronunciation, though they are heavily Japanized.
3. The Yakudoku Method also gives the student useful training for university entrance examinations in which word-for-word translations are so often required. However, the existence of those translation questions in the exams should be viewed as a result of the prevalence of yakudoku rather than as a cause of it.

References

Deyama, K. (1988). Kankoku no eigojijo. *The Chunichi Shimbun*, Jan. 20, evening edition.

Henrichsen, L. E. (1988). Diffusion of innovations in TESOL: ELEC's campaign in Japan. Paper presented at the 22nd Annual TESOL Convention, March 12, Hyatt Regency Chicago.

Hino, N. (1982). Yakudoku: The Japanese approach to foreign language study. *Working papers 1* (2), 45-53 (Department of English as a Second Language, University of Hawaii at Manoa). Also in *The Journal of the International College of Commerce and Economics, Department of Commerce*, No. 32, 85-90, 1985.

Hino, N. (1987a). *TOEFL de 650-ten: Watashi no eigoshugyo.* Tokyo: Nanundo.

Hino, N. (1987b). Eigo wo dou manabuka. *Kokusai Keizajin*, 55, 20-23.

Hiraizumi, W., & Watanabe, S. (1975). *Eigokyoiku daironso.* Tokyo: Bungeishunjo.

Ito, K. (1978). Traditional methods and new methods: A study on the methods suited for the Japanese. In I. Koike, M. Matsuyama, Y. Igarashi, & K. Suzuki (eds.), *The teaching of English in Japan* (pp. 204-219). Tokyo: Eichosha.

Ito, K. (1984). *Eigokyojuho no subete.* Tokyo: Taishukan.

Kakita, N. (1978). Yakudokushiki kyojuho. In *Eigokyujuho kakuron* (pp. 116-133). Tokyo: Kenkyusha.

Kanno, H. (1986). Nihongo to chosengo no kankei saguru. Interview by *The Asahi Shimbun*, Sept. 6, evening edition.

Kasajima, J. (1987). *Eibun sokudokusokkaiho.* Tokyo: Eikyo.

Kawasumi, T. (1975). Yakudoku no rekishi. *The English Teacher's Magazine*, July special issue, 14-19.

Kawasumi, T., with Suzuki, T. (1978). *Eigokyoiku ronsoshi.* Tokyo: Taishukan.

Koike, I., *et al.* (1983). *General survey of English language teaching at colleges and universities in Japan: Teachers' view.* Tokyo: Research Group for College English Teaching in Japan.

Koike, I., *et al.* (1985). *General survey of English language teaching at colleges and universities in Japan: Students' view.* Tokyo: Research Group for College English Teaching in Japan.

Kunihiro, M. (1970). *Eigo no hanashikata.* Tokyo: Simul Press.

Matsumoto, T. (1965). *Eigo no atarashii manabikata.* Tokyo: Kodansha.

Mombusho. (1978a). *Chugakko shidosho: Gaikokugohen.* Tokyo: Kairyudo.

Mombusho. (1978b). *Chugakko shidosho: Kokugohen.* Tokyo: Shoseki.

Mombusho. (1979a). *Kotogakko gakushushidoyoryo kaisetsu: Gaikokugohen-eigohen.* Tokyo: Hitotsubashi Shuppan.

Mombusho. (1979b). *Kotogakko gakushushidoyoryo kaisetsu: Kokugohen.* Tokyo: Gyosei.

Osawa, S., Ando, S., Kuroda, K., & Narita, Y. (1978). *More successful teaching of English.* Tokyo: Nanundo.

Ozeki, A., Takanashi, M., & Takahashi, M. (1983). *Eigoka kyoikuho.* Tokyo: Kinseido.

Sagawa, K., & Furuya, C. (1984). *Surasura eigo sokudokujutsu.* Tokyo: Sogensha.

Sogo, M. (1970). *Zusetsu Nihon no yogaku.* Tokyo: Tsukiji Shokan.

Suzuki, N. (1975). *Chugokugo to kambun.* Tokyo: Koseikan.

Tajima, K. (1978). The Grammar-Translation Method: Its historical and social background. In I. Koike, M. Matsuyama, Y. Igarashi, & K. Suzuki (eds.), *The teaching of English in Japan* (pp. 220-227). Tokyo: Eichosha.

Tazaki, K. (1978). *Theories of English teaching.* Tokyo: Taishukan.

Ueda, A. (1979). Chokudoku chokkai. In *Yomu eigo* (pp. 78-103). Tokyo: Kenkyusha.

BREAKING RULES

NEW

ALTERNATIVES IN LANGUAGE TEACHING

John F Fanselow

A unique methods book designed specially for foreign language teachers.

The author disputes the notion that teaching is a process that defies precise, rational description. By inviting the reader to classify communications in a variety of classroom as well as non-teaching settings, Dr. Fanselow presents a means for discovering rules of teacher-student interaction.

Provided with a specific, non-judgmental system of analysis, teachers can see patterns of communication they and others use, and can select alternatives they do not employ, thereby "breaking the rules" that control their teaching.

For further information
please contact Heather Saunders or
Mike Thompson at Longman ELT,
Longman Penguin Japan Company Ltd,
Yamaguchi Building,
2-12-9 Kanda Jimbocho,
Chiyoda-ku, Tokyo 101
Tel. (03) 265 7627

Longman

JALT Journal, Vol. 10, Nos. 1 & 2

FACTORS INFLUENCING THE ACQUISITION OF SECOND LANGUAGE PHONOLOGICAL COMPETENCE: CHILDREN VERSUS ADULTS

Yasushi Sekiya

Abstract

This paper examines factors affecting the acquisition of second language (L2) phonology in light of child-adult differences. Studies have shown that children achieve better L2 pronunciation than adults in naturalistic environments. Several arguments have been proposed concerning children's superiority over adults in L2 phonological acquisition. They are: (1) the neurological argument; (2) the habit formation argument; (3) the socio-affective argument; and (4) the input argument. This paper reviews each argument critically. At present there seems to be less evidence for the neurological argument and more evidence for the habit formation argument, the socio-affective argument, and the input argument. The interaction of the last three factors seems to provide a fairly adequate account of why children are superior to adults and why some adults are better than others in L2 phonological acquisition.

Linguists and lay people alike have debated whether children possess a special talent for acquiring second language L2 phonology.[1] They ask: Are children actually better at acquiring the sounds of an L2 than adults? If so, what factors explain children's superiority over adults in the acquisition of L2 phonology? The purpose of this paper is to examine these issues by referring to the literature on this subject. First, the question of whether or not differences exist between children and adults in their ability to acquire L2 phonology will be discussed. This will be followed by a discussion of factors affecting L2 phonological development.

Yasushi Sekiya teaches English at Kanda University of International Studies. He has an M.A. in linguistics from Sophia University and an M.Ed. in TESOL from Teachers College, Columbia University. He is completing his dissertation in the area of acquisition of second language phonology at Teachers College, Columbia University.

Child-Adult Differences

In the past two decades a number of studies have been conducted to examine the issue of whether children are superior to adults in acquiring the sounds of an L2 (Asher & Garcia, 1969; Olson & Samuels, 1973; Seliger, Krashen, & Ladefoged, 1975; Oyama, 1976; Snow & Hoefnagel-Hohle, 1977). (For a compendium of articles on child-adult differences in second language acquisition, see Krashen, Scarcella, & Long, 1982.) Many of them have been concerned with the critical period for language learning (from approximately age two to puberty), during which complete acquisition is possible, and beyond which native-like acquisition becomes increasingly difficult (Lenneberg, 1967). Some researchers claim that the notion of the critical period is applicable mainly to the phonological aspect of L2 acquisition (Scovel, 1969; Walsh & Diller, 1981). This is because native-like syntax seems much more attainable for adult L2 learners than native-like pronunciation (Ioup & Tansomboon, 1987).

The critical period hypothesis, as it relates to the phonological aspect of L2 acquisition, has a strong version and a weak version (Neufeld, 1980). The strong version does not allow for exceptions: No one after the age of puberty can acquire a native-like accent in an L2 (e.g. Scovel, 1969). The weak version allows for exceptions: Although most adults have difficulty acquiring native-like pronunciation, some do achieve it (e.g. Neufeld, 1980; Seliger, 1981).

The following three studies claim to have presented evidence against the argument that children are better acquirers of L2 phonology than adults. Olson and Samuels (1973) examined the pronunciation accuracy in German of three groups of 20 American subjects — elementary, junior high school, and college students with no prior knowledge of German. The groups had 10 drill sessions, each lasting 15 to 25 minutes, over a period of two weeks. Drills were conducted with language tapes in a language laboratory. Students were asked to mimic to the best of their ability words and short sentences containing 33 target phonemes. The results of the college and junior high school groups were significantly higher than those of the elementary school children. Olson and Samuels interpreted these results as indicating that adults are actually superior to children in L2 pronunciation. Thus, they claim that the critical period hypothesis is untenable.

mimic

A similar study was conducted by Snow and Hoefnagel-Hohle

(1977). The researchers investigated the Dutch pronunciation of 136 child and adult English-speaking subjects. The subjects were placed into 11 age groups ranging from 5-year-olds to 21-30 year-olds. First, subjects heard the native speaker's pronunciation of five different Dutch words containing nine target sounds a total of 20 times. Then, Repetition they were asked to repeat after the native speaker's pronunciation. The results showed that in general the older groups were better than the younger groups in pronunciation accuracy. The authors interpreted these findings as evidence against the critical period hypothesis.

Both Olson and Samuel's (1973) and Snow and Hoefnagel-Hohle's (1977) conclusions about the critical period hypothesis are misleading, however. Neither of the results attained in their studies necessarily reflects the subjects' ability to acquire L2 sounds in "naturalistic" settings[2] in the long term. The tasks used in both studies tested the ability to mimic L2 sounds in an artificial language laboratory situation. This ability seems to be closely related to the subject's concentration and length of memory span. Since children possess shorter attention and memory spans than adults, the former are at a disadvantage over the latter in these tasks. Thus, the most we could say about these two studies is that adults are superior to children in mimicking L2 sounds under language laboratory conditions at the initial phase of their learning.

In addition to their laboratory study, Snow and Hoefnagel-Hohle (1977) conducted a study of children and adults acquiring an L2 in naturalistic settings. They investigated the acquisition of Dutch by 47 English speakers in Holland. The subjects were divided into five age groups ranging from 3- to 5-year-olds to adults. They were tested three times: First within six weeks of their being exposed to a Dutch language environment, the second and third times at four- to five-month intervals. The pronunciation test consisted of 80 words which were elicited by imitation of a native speaker and by means of pictures. The results indicated that at the initial stages of acquisition, adults were superior to children in pronunciation. Based on these results, the researchers concluded that the critical period hypothesis for pronunciation could not be supported.

This conclusion is, however, questionable. Other results reported in this study indicated that although the older subjects had an initial advantage in pronunciation, they were gradually overtaken by the younger subjects. Since the critical period hypothesis concerns the eventual attainment of native-like proficiency, the findings of this study could be taken as support for, rather than refutation of, the hypothesis.

Thus, none of the above three studies invalidated the critical period hypothesis. Unlike these last three studies, the next three studies were concerned with *long-term, naturalistic* L2 acquisition. Interestingly, they all suggest that children are superior to adults in L2 phonological acquisition.

Asher and Garcia (1969) examined the relationship between the age of arrival in an L2 environment and pronunciation accuracy in 71 Cuban immigrants. Most of them had been in the United States more than five years. The subjects were divided into three groups according to their age of arrival in the United States: (1) 1-6 years old; (2) 7-12 years old; and (3) 13-19 years old. The subjects and 30 native speakers used as controls read four sentences in English into a tape recorder after rehearsing the material. Nineteen American high school students then rated each subject with respect to native-like accent, using a 4-point scale ranging from "native speaker" to "definite foreign accent."

The results showed that the younger the subject was upon arrival the better his pronunciation was likely to be. The first group, whose age of arrival was between 1 and 6, had the best chance of acquiring a near-native accent. The second group, who had arrived between ages 7 and 12, had a 50-50 chance of acquiring a near-native accent. The third group, those who had arrived at the age of 13 or older, had only a small chance of acquiring a near-native accent.

A study conducted by Oyama (1976) confirms Asher and Garcia's conclusion. Her subjects were 60 male, Italian-born immigrants who had been in the United States for a minimum of five years. They were placed in three different groups depending on their age of arrival: (1) 6-10 years old; (2) 11-15 years old; and (3) 16-20 years old. Data were collected in two ways: First by asking the subjects to read a short passage, next by asking them to describe a frightening experience in their lives. The pronunciation of each subject was then rated in terms of a 5-point scale from "no foreign accent" to "heavy accent" by two native speaker judges. The results indicated again that the younger the subject was when entering the United States, the more native-like his pronunciation tended to be.

Further support for the claim "the younger the better" is provided by Seliger, Krashen, and Ladefoged (1975). The researchers conducted a survey of 394 immigrants to the United States and Israel. The subjects were divided into three groups according to their age of arrival: (1) 9 years and younger; (2) 10-15 years; and (3) 16 years and older. They were asked to evaluate their own pronunciation and to report whether they thought other Americans or other Israelis considered them as native speakers of English or Hebrew. The researchers found that the younger the subjects were upon arrival, the more likely they were to report that others judged their pronunciation to be native-like. Since their data are based on self-report of others' informal evaluations, we have to accept these results with some reservations.

Thus, the three studies which investigated the long-term effect of age of arrival on pronunciation all suggest that children are superior to adults in the long run in naturalistic L2 phonological acquisition.

What do these studies say about the strong version of the critical period hypothesis? Both Asher and Garcia's (1969) and Oyama's (1976) studies give support to the strong version in that no subject who came to the U.S. after puberty was judged to have a native accent. However, the study done by Seliger et al. (1975) does present some evidence against the strong version. Twelve of those in the oldest arrival group (16 and older) reported that they had native-like accents. Again, however, since the data were based on self report, we have to consider these results with some reservations.

Neufeld (1980) also presented evidence suggesting that there are people who have learned a second language as adults and can pass as native speakers of that language. Seven Anglophones who learned French as adults were selected on the basis of an impressionistic judgment that they could pass as native speakers of French in casual conversation with native speakers. For this study, the subjects read a short passage into a tape recorder after having had as many practice sessions as they wanted. The researchers found that five of these subjects were consistently evaluated as native Francophones by French-speaking judges.

As is known from studies on sociolinguistic variation of L2 phonology (Dickerson, 1975; Beebe, 1980), the reading of words or passages usually (though not in all cases) produces more native-like variants than conversation. It is regrettable that Neufeld did not use conversation data for evaluation of pronunciation accuracy in conversation.

Thus, although there may be weaknesses in the nature of the data, the studies conducted by Seliger et al. (1975) and Neufeld (1980) provide some evidence for the rejection of the strong version of the critical period hypothesis.

To summarize, children have been shown to be superior to adults in naturalistic L2 phonological acquisition over the long run. Although claims have been made that there are adults who acquire native-like accents in an L2, these adults are still considered to be exceptions. What we need to know now is what factors account for children's superiority in L2 phonological acquisition over adults. This issue will be treated in the next section.

Factors Affecting Pronunciation Development

A controversy exists over why children are better acquirers of L2 pronunciation than adults in naturalistic L2 environments and why some adults are better than others in pronunciation accuracy. There are roughly four kinds of arguments to explain the reasons: (1) the neurological argument; (2) the habit formation argument; (3) the socio-affective argument; and (4) the input argument. Each of the arguments will be reviewed and discussed in the following section.

The Neurological Argument

The neurological argument basically states that the brain loses its capacity to acquire the sounds of an L2 due to lateralization (the localization of language functions in the left hemisphere) and the resultant loss of brain plasticity around the age of puberty (Scovel, 1969). Originally the neurological argument, which actually concerned all aspects of language acquisition, was proposed by Penfield and Roberts (1959), and then reinforced by Lenneberg (1967). Their claim for the completion of lateralization around puberty is based on clinical data from the recovery patterns of language function in children and adults after brain damage.

Genesee (1988) points out a number of conceptual weaknesses with the neurological argument. First, the evidence cited by Penfield and Roberts (1959) and Lenneberg (1967) concerns first language (L1) competence and, therefore, does not necessarily apply to L2 competence. Second, their evidence concerns adults with brain damage. Therefore, it does not necessarily follow that adults with normal neurological systems will experience difficulty learning an L2.

Apart from these conceptual weaknesses, however, there is still a good deal of controversy as to when lateralization is completed. By reanalyzing the data used by Lenneberg (1967) and additionally conducting dichotic listening tests, Krashen (1973) claimed that the completion of lateralization occurs by the age of 5. Furthermore, Krashen (1975) cites some researchers who argue that lateralization is basically complete at birth. If one of these claims turns out to be correct, the difficulty adults have with L2 pronunciation cannot be attributed to lateralization of the brain.

Other researchers maintain that the process of lateralization may start very early in life, but is not completed until puberty or even later in life in some people (Levy, 1974; Seliger, 1981). Based on Levy's (1974) clinical data, Seliger (1981) estimates that about 36% of a normal population still retains brain plasticity after puberty. Seliger attempts to explain the so-called exceptions to the critical period hypothesis (those people who learned an L2 as adults and have mastered native-like accent). He claims that although 36% of adults have the biological potential to acquire a native-like accent, only a small percentage of these fulfills such potential due to other inhibiting forces such as cognitive, socio-affective, and environmental factors.

Even if Seliger's (1981) conclusion regarding the termination of lateralization is correct, the relationship between lateralization and accentless speech is still highly speculative. It is not clear why lateralization of language function would mainly affect the phonological aspect of second language acquisition. Since L2 phonological acquisition involves perceptual and psychomotor skills, researchers in this field should study the neurological mechanisms related to these skills specifically.

Walsh and Diller (1981) attempted to provide a neurological explanation for the discrepancy between L2 learners' phonological ability and syntactic or lexical ability. They claim that the reason why foreign accents are difficult to overcome after childhood is that pronunciation is a lower-order process which is "dependent on the early maturing and less adaptive macroneural circuits" (p. 18). In other words, once these innate macroneural circuits, by which newborn infants are capable of detecting certain phonetic features (Elmas, 1974), are imprinted with patterns of pronunciation early in life, they do not change with experience. On the other hand, lexicon and syntax are higher-order processes which are more dependent on the late maturing neural circuits. This is why adults can learn the grammar and lexicon of an L2 many times faster than children.

Flege (1987) criticizes Walsh and Diller's (1981) argument. He maintains that their hypothesis with regard to the role of the feature detecting neural circuits would be accepted by few researchers of speech perception and is inconsistent with research results concerning the perception of L2 phonetic contrasts by adult L2 learners. In support of his criticism Flege cites studies by MacKain et al. (1981) and Williams (1980) on the perception of L2 phonetic contrasts by L2 learners. Both studies indicate that as a function of exposure to an L2, learners' perceptions of some L2 phonetic contrasts become increasingly more native-like. For example, MacKain et al. found that adult Japanese learners of English with sufficient exposure to conversational native speaker English can discriminate the members of a synthetic /r/ and /l/ continuum just as native speakers of English. Flege concludes that "there does not seem to be evidence for a discontinuity in neural development that could be reasonably regarded as coinciding with a clear change in speech-learning abilities" (p. 165).

To summarize this section, based on existing evidence no definite neurological claim can be made with regard to children's superiority over adults in L2 phonological acquisition.

The Habit Formation Argument

This argument maintains that L1 habits interfere with the acquisition of L2 habits in perception and production (Lado, 1957). Politzer (1970) suggested that the interference from L1 habits is great in adults, small in children because L1 habits are not as strongly established in children as in adults. Physiologically speaking, this may mean that the perceptual and psychomotor mechanisms of children are not constrained by L1 as much as those of adults.

More recently, Flege (1987) proposed a similar hypothesis with regard to the perception of L2 sounds by children and adults:

> Children pronounce a L2 better than adults because they tend to process speech in an 'auditory' rather than a 'phonetic' mode more often, or to a greater extent, than adults, and that this enables them to develop more accurate perceptual 'targets' for L2 sounds. An auditory mode of processing makes use of the psychoacoustic capabilities with which all individuals are endowed. A phonetic mode of processing, on

> the other hand, imposes on sensory input patterns of perceptual processing that have been shaped by previous linguistic experience. (p. 172)

Although the habit formation theory was rejected as a theory of language acquisition (Chomsky, 1959; Dulay & Burt, 1975), many empirical studies do suggest that L1 transfer plays a major role in adult L2 phonology (e.g., Carroll & Sapon, 1958; Lotz et al., 1960; Scholes, 1968; Miyawaki et al., 1975; Gass, 1984 for perception; Briere, 1966; Johansson, 1973; Flege, 1980; Tarone, 1980; Broselow, 1984; Ioup, 1985 for production. For a compendium of articles on interlanguage phonology, see Ioup & Weinberger, 1987.). Some researchers even suggest that L1 transfer is prevalent only at the phonological level, but not at the syntactic level (e.g., Ioup, 1985).

The habit formation argument as proposed here does not deny the existence of the creative construction mechanism in the acquisition of L2 phonology, however. In fact, there is substantial evidence to show that the learner's interlanguage phonology is evolving toward the target language phonology unless it is 'fossilized' (Dickerson, 1975; Flege, 1980; Gass, 1984). Furthermore, developmental processes have been reported in the acquisition of L2 phonology (for a review of such processes, see Tarone, 1978, and Sekiya, 1984). Therefore, L1 transfer here should be regarded as an important part of the L2 phonological acquisition process. Furthermore, L1 transfer phenomena includes not only mechanical transfer of exact L1 variants, as conceived by the original contrastive analysis proponents such as Lado (1957), but also other indirect cross-linguistic influence (Flege, 1980).

Although there is a great deal of impressionistic opinion that children exhibit less interference than adults in L2 phonology, there has been little empirical research measuring the actual extent to which L1 transfer plays a role in children's acquisition of L2 phonology (see Wode, 1980). So the claim that children exhibit much less L1 influence in the acquisition of L2 phonology than adults awaits further verification.

Unlike the neurological argument, which considers puberty to be a sudden turning point for L2 phonological acquisition, the habit formation argument does not specify any such age, but suggests that the younger one learns an L2 the better. Therefore, the habit formation argument can explain not only why children are better than adults but also why younger children are better than older children in the acquisition of native-like speech (Asher & Garcia, 1969; Oyama, 1976; Seliger, Krashen, & Ladefoged, 1975).

Furthermore, unlike the neurological argument, the habit formation argument does not imply the ultimate loss of the ability to acquire accurate L2 sounds. Neufeld (1978) demonstrated that with proper training adults are biologically capable of learning new L2 sounds. In this experiment, 20 English-speaking college students were given a special intensive training in the pronunciation of 100 stock phrases of Japanese, Chinese, and Eskimo. The results of this experiment indicated that about 50% of his subjects were judged to have native-like accent. Further research needs to be conducted to see if their native-like pronunciation would carry over into actual communication. However, Neufeld's results demonstrated that many adults have not lost their *biological* capability to perform new articulatory behaviors of L2 sounds, at least under rigorous conditions.

One important point about Neufeld's (1978) technique of pronunciation teaching is that the subjects did a great amount of listening, focusing on rhythm, pitch, and intonation contours of the utterances before they were allowed to imitate them. Neufeld maintains that the learner's speech perception is influenced by his L1 phonology. Once inaccurate acoustic images of L2 pronunciation are formed due to L1 interference and are fixed in the learner's mind, they also affect the articulation of these sounds. Therefore, it is important for the learner to establish accurate sound images of L2 sounds before attempting to produce them. Neufeld's argument is interesting in that children in a naturalistic L2 environment are also observed to do a great amount of attentive listening before they attempt to speak an L2 (Dulay, Burt, & Krashen, 1982).

To summarize, the habit formation argument is plausible in that it can account for child-adult differences in terms of the amount of L1 interference. However, the claim that children experience much less interference from their L1 phonology than adults should be investigated empirically. Furthermore, it has been suggested that with proper training adults can acquire native-like pronunciation.

The Socio-Affective Argument

The socio-affective argument suggests that the learner's socio-affective traits such as attitudes, motivation, cultural identity, and empathy will influence his achievement in L2 phonological acquisition. Krashen (1982) suggests that socio-affective variables influence progress in second language acquisition in at least two ways. First, learners with positive socio-affective states will attempt to communicate more with speakers of the target language and therefore receive more input than learners without such socio-affective states. (This point will be discussed later in the section on the input argument.) Second, L2 learners with positive socio-affective states will be more open to L2 input. That is, given the same amount of input, learners with positive socio-affective states will acquire more input than learners without such socio-affective states. In the light of this second point the effects of socio-affective variables on the acquisition of L2 phonology will now be discussed.

With the onset of the formal operational stage[3] around puberty, children become increasingly conscious of themselves (Elkind, 1970). This also relates to increasing awareness of their cultural/ethnic identity. They will develop a stronger sense of cultural/ethnic allegiance or negative social attitudes towards one language/culture or another[4] (Schumann, 1975; Brown, 1980). As a result, it will be harder for them to adopt a new cultural/ethnic identity. These changes in the learner's socio-affective states seem to have a great consequence in L2 learning, including L2 phonological acquisition.

There is substantial evidence to suggest that pronunciation is an important aspect of cultural/ethnic identity (Bourhis & Giles, 1977; Beebe, 1977; Beebe & Giles, 1984). If children in an L2 environment experience relative ease adopting a new cultural identity, we should expect them to adopt L2 pronunciation relatively easily too. On the other hand, many adults might have difficulty assimilating new pronunciation behaviors because of a psychological block against adopting a new cultural identity.

However, some adults have integrative motivation to learn a second language, that is motivation to achieve L2 proficiency in order to participate in the target culture or to be a member of the target society (Gardner & Lambert, 1972). These adults may be willing to adopt new cultural norms, and may thus try to improve their L2 pronunciation. Seliger et al.'s study (1975), previously described, presented some evidence to support the claimed relationship between integrative motivation and pronunciation achievement. Among the subjects in the 10 to 15 age group, those with no accent tended to consider themselves "more American" than those with an accent. These results concur with Gardner and Lambert's conclusion that learners with integrative motivation develop better oral skills.

Another important observation is that children are under more pressure to conform to the norms of their peers, including norms of pronunciation (Tarone, 1978; Peck, 1978; Brown, 1980; Snow & Hoefnagel-Hohle, 1977). Children even mock the child learner's accent directly (Peck, 1978). Thus, in order to be accepted into their peer group, child L2 learners try to adopt the same accent as their peer group. On the other hand, adults are not only less susceptible to peer pressure to conform but are often willing to retain their foreign accents in order to maintain their cultural/ethnic identity (Snow & Hoefnagel-Hohle, 1977; Beebe & Giles, 1984).

Adolescents in an L2 environment are also under pressure to conform to the norms of their peers (Oyama, 1976). However, adolescent L2 learners are considered to have a disadvantage over child L2 learners in that having already developed a cultural/ethnic identity, adolescents have more "perceptive filters to readjust" (Brown, 1980, p. 139). From this argument, we should expect children to surpass adolescents, and adolescents to surpass adults in L2 phonological attainment in naturalistic environments. The data reported by Oyama (1976) and Seliger et al. (1975) support this hypothesis.

It should be noted that in a foreign language classroom, where the learners are all non-native speakers of an L2 from the same L1 background, peer pressure may work against L2 pronunciation development. The learners may make little effort to master L2 pronunciation for fear of sounding "foreign" or different from others in the class, thus retaining their L1 accent in their L2 (Stevick, 1976; Hildebrandt & Giles, 1981).

With regard to the relationship between identity and pronunciation, Guiora et al. (1972a, 1972b, 1980) proposed an interesting argument. According to Guiora et al., pronunciation is the most important contribution of "language ego" (the identity a person develops in relation to the language he speaks) to a person's identity. They suggested that children's flexible ego boundaries allow them to identify or empathize with the speakers of a new language, and as a result, children in an L2 environment assimilate native-like pronunciation of the language relatively easily. However, this is not usually the case with many adults, who have already established firm ego boundaries.

Nonetheless, some adults are considered to have higher empathy levels than others. Guiora et al. (1972a, 1972b, 1980) suggest that the higher the person's empathy level is, the better chance he has of achieving good L2 pronunciation. Guiora et al. attempted to demonstrate the correlation between empathy and L2 pronunciation ability experimentally. In their first experiment, Guiora et al. (1972a) attempted to increase the empathy levels of subjects through the use of alcohol. They found that the pronunciation of subjects given a small amount of alcohol was significantly better than that of a control group. Brown (1980) criticized this experiment by arguing that muscle relaxation caused by alcohol may be a more important factor in accounting for the superior pronunciation performance of the subjects than the affective effect. To address this criticism, Guiora et al. (1980) replicated the alcohol study using valium instead of alcohol as a means of increasing subjects' empathy levels. The results indicated a relationship between empathy and pronunciation, according to researchers.

Hill (1970) also supports the socio-affective argument based on some anthropological evidence. Challenging Scovel's (1969) neurological argument for the critical period hypothesis, Hill introduces Sorenson's (1967) report that in some non-Western societies, where multilingualism is highly valued, adults acquire native-like fluency in second languages. Hill interpreted this report as indicating that lack of empathy with the speakers of other languages on the part of adults is not universal, but culturally-determined.

One serious weakness with Hill's argument is that it is based on second-hand anecdotal evidence. There is no empirical evidence to show that those adult L2 learners whom Sorenson reports to have achieved native-like proficiency actually speak the language with native-like accent. However, if further research shows that this is the case, it would be strong evidence against the neurological argument for the critical period hypothesis and strong support for the socio-affective argument.

To summarize this section, the socio-affective argument is certainly appealing in that it can explain both child-adult differences and individual variation among adults in L2 phonological attainment. At present, however, much of the evidence is impressionistic. More research is needed.

The Input Argument

The input argument is not equivalent to Krashen's (1985) Input Hypothesis. The input argument states that the quantity, the quality, and the conditions of input are important in determining the degree of success in achievement of native-like accent.

It has been mentioned previously that in general, children in an L2 environment are highly motivated to assimilate with their peers, while adults are not. This difference in integrative motivation will naturally relate to the amount of contact with speakers of the target language. This, in turn, will relate to the amount of native speaker input and the intensiveness of the input they will receive. Those with integrative motivation, children or adults, will have more contact with target language speakers; thus they will receive more intensive native speaker input than those without integrative motivation (Gardner & Lambert, 1972; Seliger et al., 1975; Oller, 1977; Beebe & Giles, 1984). Snow and Hoefnagel-Hohle (1978) report that English-speaking children learning Dutch in Holland received more L2 input than adults.

In addition to the amount of native speaker input, how that input is received by L2 learners also seems to affect the nature of L2 phonological development. As mentioned in the discussion of the habit formation argument, many children are reported to go through a "silent period," during which they do a great deal of intensive listening. This period will help them to form accurate acoustic images of L2 segmental and suprasegmental sound features, which is subsequently crucial for their accurate production (Neufeld, 1978). On the other hand, adults, pressured to communicate from the start, often have to produce L2 sounds before they have formed accurate acoustic images. Once inaccurate acoustic images of the L2 sounds become fixed in the learner's mind, they affect production (Neufeld, 1978). Interestingly, the Vaupes River Indians in Brazil, who are reported to acquire a native-like command of an L2 as adults, also engage in attentive listening before they try to produce (Sorenson, 1967).

Related to the silent period argument is the claim that children in a naturalistic L2 environment depend on the 'ear' more than adults for pronunciation. Adults are more likely to depend not only on native speaker input but also on other sources for pronunciation, such as orthography and articulatory explanation of L2 sounds.[5] These other sources of pronunciation may interfere with the learner's formation of accurate acoustic images of L2 sounds, and thus their production. Sekiya (forthcoming) provides some data to show children's dependence on the acoustic images of L2 sounds for pronunciation. In a study on the acquisition of English pronunciation by 80 Japanese children in the United States, the researcher reports many instances of the substitution of /f/ for /θ/ among the subjects. This substitution is attributable to the acoustic similarity of /θ/ and /f/ (Delattre, Liberman, & Cooper, 1962), and is often found among children acquiring English as L1 (Edwards & Shriberg, 1983). This kind of substitution is not reported among Japanese adult learners (Kohmoto, 1975).

To summarize this section, children seem to have an advantage over adults in terms of the native speaker input that they receive in an L2 environment. The amount of input, the quality of input, and how children receive input all seem to be favorable for their acquisition of L2 phonology.

Factors Related to Adults' Varying Degrees of Success in L2 Pronunciation Achievement

In the previous sections I have examined factors influencing L2 pronunciation development in terms of child-adult differences. The following studies by Suter (1976) and Purcell and Suter (1980) attempt to account for adults' individual variation in their ability to acquire L2 phonology. Suter (1976) studied the correlations between English pronunciation accuracy scores and 19 variables for 60 non-native speakers of English from four language backgrounds: Persian, Arabic, Japanese, and Thai. The results of Purcell and Suter's (1980) reanalysis of Suter's original data indicated that the following four variables were the most significant predictors of pronunciation accuracy: (1) L1 background; (2) aptitude for oral mimicry; (3) length of residency in the L2 country; and (4) strength of concern for pronunciation accuracy. It should be noted that age of arrival was not included among the four best predictors.

In interpreting these results, several important points are implied. First of all, the finding that age was not such a significant predictor as these four variables may seem contradictory in light of the evidence presented in the first section of this paper. But actually it is not contradictory. This result seems to be attributable to the fact that none of the subjects had arrived in the U.S. before pubeerty. If prepubertyarrival subjects had been included, age would likely have been among the significant predictors.

Second, two possible interpretations of the most important predictor, L1 background, were suggested by Suter (1976): (1) Phonological and phonetic differences between L1 and L2 may have affected the learner's pronunciation accuracy, or (2) cultural or personality traits of each language group may have affected their pronunciation accuracy. The first interpretation is related to the habit formation argument. The second is related to the socio-affective argument, more specifically Hill's (1970) claim that different cultures have different attitudes towards learning other languages, resulting in different degrees of L2 proficiency.

Conclusion

Accumulated evidence supports the popular claim that children are superior to adults in the long run in the acquisition of L2 phonology in naturalistic settings. Several possible explanations for this were examined in this paper. At present, the neurological argument does not have as much empirical support as it was once claimed to have. However, it is possible that there is still some neurological component in children's superiority over adults. The habit formation argument, the socio-affective argument, and the input argument also seem to be plausible as explanations for the children's advantage, and these arguments have more support. They are not necessarily mutually exclusive. In fact, it is likely that these three factors affect the acquisition of L2 phonology in one way or another. The interaction of these factors seems to provide a fairly adequate account of why children are superior to adults and why some adults are better than others in L2 phonological acquisition. Future research should be conducted to provide further empirical support for each argument and to examine to what extent each factor and the interaction of the factors influence the acquisition of L2 phonology.

[This article is a revised version of a paper presented at the 13th JALT International Conference in Tokyo, November 21-23, 1987. The author thanks Leslie Beebe, Lynn Broquist, Robert DeSilva, Monica Hamolsky, Thurston Womack, and Walt Wolfram for their valuable comments on an earlier version of this paper.]

Notes

1. L2 phonology includes not only L2 segmental features but also L2 suprasegmental features.
2. The term "naturalistic" is used to refer to an L2 acquisition environment in which an L2 is used as a medium of communication. It is in contrast with a foreign language classroom environment or an experimental learning environment.
3. The formal operational stage is a Piagetian concept of a stage of children's development. With the onset of the formal operational stage, children can deal with abstract concepts and make hypotheses, inferences, and deductions.
4. Children are reported to begin to acquire certain attitudes toward one language/culture or another as they reach school age (Brown, 1980).
5. This change in the learner's learning strategies may be related to the cognitive change people go through during the stage of formal operations around puberty (see Note 3). Due to this cognitive change, it is now possible for L2 learners to 'learn' the rules of the language *consciously* in addition to 'acquiring' them *subconsciously* in a manner similar to children (Krashen, 1981).

References

Asher, J., & Garcia, R. (1969). The optimal age to learn a foreign language. *Modern Language Journal 53*, 334-342.

Beebe, L. (1977). The influence of the listener on code- switching. *Language Learning, 27*, 331-339.

Beebe, L. (1980). Sociolinguistic variation and style shifting in second language acquisition. *Language Learning, 30*, 433-448.

Beebe, L. M., & Giles, H. (1984). Speech accommodation theories: A discussion in terms of second language acquisition. *International Journal of the Sociology of Language, 46*, 5-32.

Bourhis, R. Y., & Giles, H. (1977) The language of intergroup distinctiveness. In H. Giles (ed.), *Language ethnicity and intergroup relations* (pp. 119-136). New York: Academic Press.

Briere, E. J. (1966): An investigation of phonological interference. *Language, 42*, 786-796.

Broselow, E. (1984). An investigation of transfer in second language phonology. *IRAL, 22*, 253-269.

Brown, H. D. (1980). *Principles of language learning and teaching*. Englewood Cliffs, NJ: Prentice-Hall

Carroll, J., & Sapon, S. (1958). Discriminative perception of speech sounds as a function of native language. *General Linguistics, 3*, 62-71.

Chomsky, N. (1959). Review of *Verbal behavior* by B.F. Skinner. *Language, 35*, 26-58.

Delattre, P. C., Liberman, A. M., & Cooper, F. S. (1962). Formant transitions and loci as acoustic correlates of place of articulation in American fricatives. *Studia Linguistica, 16*, 101-121.

Dickerson, L. (1975). The learner's interlanguage as a system of variable rules. *TESOL Quarterly, 9*, 401-408.

Dulay, H., & Burt, M. (1975). Creative construction in second language learning and teaching. In H. Dulay & M. Burt (eds.), *New directions in second language learning, teaching and bilingual education* (pp. 21-32). Washington, DC: TESOL.

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

Edwards, M., & Shriberg, L. (1983). *Phonology: Applications in communicative disorders.* San Diego: College-Hill Press.

Eimas, P. D. (1974). Auditory and linguistic processing for cues for place of articulation by infants. *Perception and Psychophysics, 16,* 513-521.

Elkind, D. (1970). *Children and adolescents: Interpretative essays on Jean Piaget.* New York: Oxford University Press.

Flege, J. E. (1980). Phonetic approximation in second language acquisition. *Language Learning, 30,* 117-134.

Flege J. E. (1987). A critical period for learning to pronounce foreign languages? *Applied Linguistics, 8,* 162-177.

Gardner, R. C., & Lambert, W. E. (1972). *Attitudes and motivation in second language learning.* Rowley, MA: Newbury House.

Gass, S. (1984). Development of speech perception and speech production abilities in adult second language learners. *Applied Psycholinguistics, 5,* 51-74.

Genesee, F. (1988). Neuropsychology and second language acquisition. In L. M. Beebe (ed.), *Issues in second language multiple perspectives* (pp. 81-112). New York: Newbury House.

Guiora, A., Beit-Hallahmi, B., Brannon, R., Dull, C., & Scovel, T. (1972a). The effects of experimentally induced changes in ego states on pronunciation ability in a second language: An exploratory study. *Comprehensive Psychiatry, 13,* 421-428.

Guiora, A., Brannon, R., & Dull, C. (1972b). Empathy and second language learning. *Language Learning, 22,* 111-130.

Guiora, A., Acton, W., Erard, R., & Strickland Jr., F. (1980). The effects of benzodiazepine (valium) on permeability of language ego boundaries. *Language Learning, 30,* 351-363.

Hildebrandt, N., & Giles, H. (1980). The English language in Japan: A social psychological perspective. *JALT Journal, 2,* 63-87.

Hill, J. (1970). Foreign accents, language acquisition, and cerebral dominance revisited. *Language Learning, 20,* 237-248.

Ioup, G. (1985). Is there a structural foreign accent? A comparison of syntactic and phonological errors in second language acquisition. *Language Learning, 35,* 1-17.

Ioup, G., & Tansomboon, A. (1987). The acquisition of tone: A maturational perspective. In G. Ioup & S. H. Weinberger (eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 333-349). New York: Newbury House.

Ioup, G., & Weinberger, S. H. (eds.) (1987). *Interlanguage phonology: The acquisition of a second language sound system.* New York: Newbury House.

Johansson, F. A. (1973). *Immigrant Swedish phonology: A study in multiple contact analysis.* Lund, Sweden: CWK Gleerup.

Kohmoto, S. (1975). *New English phonology: A contrastive study of English and Japanese pronunciation.* Tokyo: Nan'un-do.

Krashen, S. D. (1975). The development of cerebral dominance and language learning: More new evidence. In D. Dato (ed.), *Georgetown University round table on language and linguistics* (pp. 209-233). Washington, DC: Georgetown University Press.

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

Krashen, S. D. (1982). Accounting for child-adult differences in second language rate and attainment. In S. D. Krashen, R. C. Scarcella, & M. H. Long (eds.), *Child-adult differences in second language acquisition* (pp. 202-226). Rowley, MA: Newbury House.

Krashen, S. D., Scarcella, R. C., & Long, M. H. (eds.) (1982). *Child-adult differences in second language acquisition.* Rowley, MA: Newbury House.

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

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers.* Ann Arbor: The University of Michigan Press.

Lenneberg, E. H. (1967). *Biological foundations of language.* New York: John Wiley & Sons.

Levy, J. (1974). Psychobiological implications of bilateral asymmetry. In S. Dimond & J. Beaumont (eds.). *Hemisphere function in the human brain.* New York: John Wiley & Sons; cited in Seliger, 1981.

Lotz, J., Abramson, I., Gertsman, I., Ingemann, J., & Nemser, W. (1960). The perception of English stops, by speakers of English, Spanish, Hungarian and Thai: A tape-cutting experiment. *Language and Speech, 3,* 71-77.

MacKain, K., Best, C., & Strange, W. (1981). Categorical perception /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics, 2,* 369-390.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics, 18,* 331-340.

Neufeld, G. G. (1978). On the acquisition of prosodic and articulatory features in adult language learning. *Canadian Modern Language Review, 34,* 163-194.

Neufeld, G. G. (1980). On the adults' ability to acquire phonology. *TESOL Quarterly, 14,* 285-298.

Olson, L., & Samuels, S. (1973). The relationship between age and accuracy of foreign language pronunciation. *Journal of Educational Research, 66,* 263-268.

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Review, 5,* 261-284.

Peck, S. (1978). Child-child discourse in second language acquisition. In E. M. Hatch (ed.), *Second language acquisition: A book of readings* (pp. 383-400). Rowley, MA: Newbury House.

Penfield, W., & Roberts, L. (1959). Speech and brain mechanisms. New York: Atheneum.

Politzer, R. L. (1970). *Foreign language learning: A linguistic introduction.* Englewood Cliffs, NJ: Prentice-Hall.

Purcell, E. T., & Suter, R. W. (1980). Predictors of pronunciation accuracy: A reexamination. *Language Learning, 30,* 271-287.

Schumann, J. (1975). Affective factors and the problem of age in second language acquisition. *Language Learning, 25,* 209-235.

Scholes, R. (1968). Phonemic interference as a perceptual phenomenon. *Language and Speech, 11,* 86-103.

Scovel, T. S. (1969). Foreign accents, language acquisition, and cerebral dominance. *Language Learning, 19,* 245-254.

Sekiya, Y. (1984). The acquisition of second language phonology: A comparison with the acquisition of first language phonology. *JACET Bulletin, 15,* 19-37.

Sekiya, Y. (forthcoming). The acquisition of English phonology by Japanese children learning English in natural settings. Unpublished doctoral dissertation, Teachers College, Columbia University, New York.

Seliger, H. W. (1981). Exceptions to critical period predictions: A sinister plot. In R.W. Andersen (ed.), *New dimensions in second language acquisition research* (pp. 47-57). Rowley, MA: Newbury House.

Seliger, H. W., Krashen, S. D., & Ladefoged, P. (1975). Maturational constraints in the acquisition of second language accent. *Language Sciences, 36*, 20-22.

Snow, C., & Hoefnagel-Hohle, M. (1977). Age differences in the pronunciation of foreign sounds. *Language and Speech, 20*, 357-365.

Sorenson, A. (1967). Multilingualism in the northwest Amazon. *American Anthropologist, 69*, 670-684.

Stevick, E. (1976). *Memory, meaning, and method.* Rowley, MA: Newbury House.

Suter, R. W. (1976). Predictors of pronunciation accuracy in second language learning. *Language Learning, 2*, 233-253.

Tarone, E. (1978). The phonology of interlanguage. In J. C. Richards (ed.), *Understanding second and foreign language learning* (pp. 15-33). Rowley, MA: Newbury House.

Tarone, E. (1980). Some influences on the syllable structure of interlanguage phonology. *IRAL, 18*, 139-152.

Walsh, T., & Diller, K. (1981). Neurolinguistic considerations on the optimal age for second language learning. In K. Diller (ed.), *Individual differences and universals in language learning aptitude* (pp. 3-21). Rowley, MA: Newbury House.

Williams, L. (1980). Phonetic variation as a function of second-language learning. In G. Yeni-Komshian, J. Kavanagh, & C. Ferguson (eds.), *Child Phonology: Vol. 2. Perception* (pp. 185-215). New York: Academic Press.

Wode, H. (1980). Phonology in L2 acquisition. In S. W. Felix (ed.), *Second language development: Trends and issues* (pp. 123-136). Tubingen, West Germany: Gunter Narr Verlag.

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 Review 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.

GET YOUR STUDENTS OUT OF THE CLASSROOM!

Put **LIVING AND WORKING IN AMERICA** into the Classroom!

VIA-USA's **LIVING AND WORKING IN AMERICA**, a new video series, is a very effective means of teaching verbal communication in English. Each of the three volumes consists of separate video sequences which give short situational dialogues as well as visual information on the way Americans interact.

Each volume consists of:

- ☐ two video cassettes with a total playing time of approximately 70 minutes.
- ☐ a student text
- ☐ a teacher's manual with clearly written notes on presentation, follow-up activities, and culture notes for each lesson.

Major areas of emphasis include the following:

- ☐ socialization skills
- ☐ using airports, hotels, taxis
- ☐ using the telephone
- ☐ running meetings
- ☐ giving job interviews
- ☐ handling emergencies
- ☐ opinionating
- ☐ problem solving
- ☐ and *much more*

FOR MORE INFORMATION CONTACT:

(In Japan)
Mitsui & Co., Ltd., Osaka Office
Industrial Film Section
3-33 Nakanoshima 2-chome
Kita-ku, Osaka 530
Attn. Mr. Hirotoshi Takeuchi
☎ (06) 226-3055

(In the United States)
VIA Press, Inc.
400 E. Evergreen Blvd., Suite 314
Vancouver, WA 98660
U.S.A.

☎ (206) 696-0900

JALT Journal, Vol. 10, Nos. 1 & 2

A REASSESSMENT OF ERROR-COUNT

Jonathan D. Picken

Abstract

Research in the field of error-count evaluation of EFL writing has not had a history of results consistently successful enough to establish the method as a full-fledged alternative to analytic or general impression marking. In this article it is suggested that this record is not so much due to an inherent weakness of error-count evaluation as such, but rather to the lack of a theoretical foundation in the methods used to date.

In order to make this point, extensive reference is made to the Dutch CITO writing proficiency test (CITO, 1984) and to a related CITO study, Melse and Verstralen (1986). Data from the latter are marshalled by the present author in making his case in support of error-count evaluation. The described evaluation procedure is used to determine the English language writing skills of test subjects, and not to correct student essays.

Error-Count Evaluation of EFL Writing and the CITO Writing Proficiency Test

Over the years, error-count (EC) methods of writing proficiency assessment — variously known as frequency-count or objective methods — have not had an easy time in establishing themselves as viable alternatives to analytic or to general-impression approaches. While coming under fire for a lack of theoretical underpinnings, EC has been hard put to come up with results consistent enough to silence its critics. The problem has been compounded by the fact that EC, concentrating as it does on errors, has found itself out of step with mainstream writing pedagogy where work on mistakes has had to take something of a back seat, and only errors that interfere with communication have been deemed worthy of attention.

Jonathan Picken has obtained Master's degrees from Groningen State University in the Netherlands, and the University of London Institute of Education. He is a full-time lecturer at Tokai University in Kanagawa prefecture, and has lectured on writing proficiency testing at two JALT conferences.

The key issue that has not been addressed so far is the question of whether EC as a method is in principle misguided. Criticism so far has concerned itself with the weaknesses of EC studies to date, without considering the broader issue of whether EC in whatever form is bound to fail. The purpose of this paper is to show that the case is far from closed, and that EC with a broader theoretical orientation has a very considerable potential that deserves to be recognized.

The Background

Jacobs, Zinkgraf, Wormuth, Hartfiel, and Hughey (1981) define EC methods as ones that

> tally or enumerate certain elements in the composition such as: the number or type of words, clauses, T-units [see Table 1], cohesive devices, misspelled words, misplaced commas, or sentence errors. (p. 29)

The key word is *tally*, for this is what distinguishes EC from general impression or analytic marking. The EC marker counts; markers using the latter methods do not.

As stated in the introduction, the validity of all this counting has been called into question, mainly because it tends to ignore what are considered major aspects of student compositions. Perkins (1983), a former EC researcher, writes one of the strongest formulations of this view, claiming that EC methods are

> of little value in assessing the underlying constructs of writing [and that] currently used objective methods do not quantify cohesion, coherence, organization, . . . idiom, diction, tone, relevance or focus — all factors which contribute to good writing. (p. 662)

Against the background of this criticism, it is unfortunate, though perhaps not surprising, that even some of the most popular EC methods have failed to distinguish consistently between students of different ability levels. A comparison of the performance of four of these measures over three studies (see Table 1) shows that not one of them succeeds in distinguishing *both* between students of roughly the same level of ability, and between students of different ability levels. Thus the same measure may differentiate well between, for example, good and poor beginners, but not between, for example, beginners and intermediate students. This is clearly unsatisfactory.

Table 1
Ability of Four EC Methods to Evaluate Performance
within levels and Distinguish *between levels*

Basis of Evaluation	Perkins (1980)	Larsen-Freeman (1983)	Homburg (1984)
length error-free T-unit	within level	only between 2 of her 4 levels	
errors per T-unit	within level		not between levels
total errors per composition	within level		not between levels
error-free T-units per composition	within level	not between levels	between levels

(Note: T-units or terminable units are defined as the shortest possible units of a passage that are grammatically allowable to be punctuated as sentences.)

Further criticism could be adduced, but we need not labor the point; with neither a claim to theoretical respectability, nor a solid record of good experimental results, the case for EC methods seems seriously flawed.

Limitations of Research to Date

The main weakness of research carried out so far has been that the theoretical justification for the EC methods used has often been questionable. Given this situation, it would have been purely coincidental if any of the measures had performed more satisfactorily than they did. There is no obvious reason, after all, to expect that a measure of writing ability based on, for example, a mechanical tallying of the number of errors in a composition is likely to perform in the same way as an analytic rating scale that takes, for example, content and organization into consideration. Research so far seems to bear out this point.

The question that has gone unanswered, however, is whether an EC method that does start from an informed perception of what matters in composition would fare as badly as its predecessors. In the course of the following discussion of the CITO writing proficiency test, I hope to provide at least a partial answer to this question.

The CITO Writing Proficiency Tests

The writing tests that will be discussed in this section were developed by the CITO, the Dutch Central Institute for Test Research, with the aim of assessing the writing proficiency of students in their final, examination year at three different types of secondary school: VWO, HAVO, and MAVO. The VWO has a six-year curriculum, and the final examination at this school admits successful students to university education. Graduates of HAVO and MAVO — the former has a curriculum of five and the latter of four years — are admitted to various other kinds of further education, but not to universities. In their examination years, VWO, HAVO, and MAVO students would typically be 18, 17, and 16 years old respectively.

The CITO tasks are highly controlled and require students to write formal or informal letters to a more or less clearly defined audience and with a purpose and content specified in detail by means of a number of sub-assignments formulated in Dutch. The tasks vary in degree of difficulty according to the kind of school for which they were designed.

The marking protocol used (see Table 2 below) is the same for all three types of school. A notable contribution to its final form came from a group of schoolteachers who cooperated with CITO in carrying out trials and providing feedback. It was essential to take these future users' views into account, because within the Dutch educational system, secondary-school teachers were under no obligation to use these particular tests for examination-year writing proficiency assessment. Consequently widespread acceptance of the tests could only be achieved by gaining a broad consensus of support among teachers. The teachers insisted on having a much more detailed and specific marking protocol[1] than CITO had originally intended (Melse, 1984, p. 358). Further support for such changes came from the test trials (Melse, 1984, p. 358).

Given the test's history, it will come as no surprise to find that the marking protocol is not the product of one single perception of writing proficiency. It does, however, take a very broad range of errors into consideration. Beyond punctuation, spelling, and grammatical errors, it looks for errors of style (excessive repetition), discourse (illogical connection; lack of clarity; too sudden change of subject) and of appropriateness (use of word or expression inappropriate to context) (see Appendix I for a clarification of these error types). Content errors are manifested as "incompleteness" or "absolutely incomplete" in cases where students have failed to carry out a sub-assignment partially or completely. The bonus/malus or penalty system allows for a limited number of points to be added or subtracted where unquantified strengths or weaknesses of compositions require this.[2]

Table 2

Summary of the CITO Marking Protocol (after CITO, 1984, p. 20)

Sign	**Meaning**	**Characteristics**	**Points**
O (O O etc.)	incompleteness	element of the sub-assignment missing	–2 (–4 etc.)
?	absolutely incomplete	assignment not carried out	–8
+ (+ +)	bonus	style or content of item observably above average	+1 (2)
– (– –)	malus	style or content of item observably below average	–1 (2)
═══	wordgroup error	network of errors such that it is difficult to decide how many errors have been made	–2
———	primary error	— removal or replacement of word required — grammatical error — word-order error — excessive repetition — illogical connection — lack of clarity — use of word or expression inappropriate to context	–1
—V—	primary error	— date partially missing — word needs to be added — too sudden change of subject	–1
—/—	secondary error	— spelling error — secondary preposition error	–1/2
X	punctuation	— punctuation, apostrophes, capital and lower-case letter errors	–1/4
	dependent error, repetition of error, punctuation error of type not to be counted		

All students start with a basic score of 40 points. After correction, points are added to and subtracted from this total. The resulting total is translated into a score on a scale of 10, using CITO score conversion tables. (See Appendix A for definition of Characteristics. See Appendix B for a sample letter marked along these lines.)

In broadness of orientation, the CITO test is indisputably superior to its predecessors, and this is reflected in the test statistics, for the reliability figures are, if not ideal, at least sufficiently strong for our present purposes. Thus if the tests are administered according to CITO instructions (students write one formal and one informal letter on separate occasions, with each letter being marked independently by two teachers), test-retest reliability scores of around .80 are achieved. (A score of 1.0 would mean complete agreement among markers.)

Test-retest reliability is a measure of the extent to which the same test administered on a different occasion produces the same result. Table 3 below shows how CITO's reliability scores improve when either the number of compositions or the number of markers or both are increased. The scores of roughly .80 referred to in the previous paragraph are displayed in the fifth column (2 compositions, 2 markers). In all cases, the reliability has been checked by giving a separate test and comparing the point result with that of the earlier test(s). Thus, the figures in column one were reached when results of a second test were used to determine the accuracy of the first evaluation.

Table 3
Test-Retest Reliability of the CITO Writing Proficiency Test

	Number of compositions					
	1			2		
	number of markers			number of markers		
	1	2	3	1	2	3
HAVO English	.51	.68	.76	.68	.81	.86
VWO English	.49	.66	.74	.66	.79	.85

(Melse & Verstralen, 1986, p. 122)

Interrater reliability, that is the extent to which different raters assign the same scores to compositions, consistently hovers around the .65 mark (Melse & Verstralen, 1986, p. 111). Intrarater reliability, which shows the degree to which a rater is consistent in assigning grades, was not calculated.

To validate the test, CITO carried out a rank-order correlation of 25 letters (12 formal, 13 informal) rated by, on the one hand, a group of 10 HAVO/VWO teachers, and on the other a group of 16 NS raters (teachers and non-teachers) who ranked the letters using general impression. The resulting Spearman's rank-correlation figures[3] are r=.92 for the informal letters, and r=.86 for the formal ones. These results show that native speakers and Dutch foreign language teachers rank writing samples similarly when the latter use the CITO method.[4]

The reliability figures compare favorably with the popular Jacobs et al. (1981) ESL Composition Profile ones. Jacobs et al. have a test-retest reliability score of .72 (p. 73) and an interrater reliability coefficient of .65 (p. 69).[5]

The significance of all these figures, within the context of this article, is not so much that they tell us how good a test the CITO's is, but rather, what they by extension suggest concerning the potential of similar EC methods in general. Earlier we saw that uni-dimensional EC measures failed to distinguish consistently between levels of proficiency. However, once we start using a measure that is multi-dimensional and to take a broad range of errors into account, the picture changes We get a test that, in terms of conventional reliability and validity falls broadly within the range of currently available analytic and general impression tests of writing proficiency. The test, in other words, that the case for EC methods of writing proficiency is still very much an one.

A Drawback of EC Marking

One important drawback of the EC method, and especially one along CITO lines, is that it is more time-consuming than either analytic or general marking, and it is an open question as to whether future improvements of EC in this respect will serve to reduce the difference significantly. At present EC would appear to be inappropriate for large-scale testing where cost-efficiency imposes stringent limitations on the time allowed for marking. EC would seem to be much more suitable for contexts such as the Dutch one where teachers, being markers of their students' own compositions, need a method that comes up with consistent and valid scores and at the same time provides students with detailed information on how their scores have been determined.

Conclusion

In the course of this article, it has been argued that up to the present the case for EC methods of grading has been poorly made. Such measures as "errors per T-unit" or "error-free T-units per composition" are not only too narrow from the theoretical point of view, but they also fall far short of a convincing performance in practice. Unfortunately it has also been assumed that, by extension, EC in general has no merit.

This conclusion is premature. The CITO writing proficiency tests show that if errors are weighted and categorized so that many relevant aspects of compositions are included in their assessment, EC can operate with the same success as general impression or analytic methods of marking. In addition it seems likely that EC marking could be significantly improved by means of further research into methods of delineating error-categories even more clearly and by looking into ways of relating these categories to aspects of writing proficiency drawn from models of communicative competence.

Notes

1. In fact CITO calls the method analytic, even though it clearly belongs to the error-count category, that is, it requires markers to count errors and deduct the resulting, weighted, total from a basic score of 40.
2. The contribution of the bonus/malus system to score variance is very small: .4% (Melse & Verstralen, 1986, p. 111).
3. CITO uses a rather more complicated statistical method to calculate the correlations. In the main text I have reported Spearman's rank correlation figures — based on my own re-analysis of the data — as readers are more likely to be familiar with this method.
4. Strictly speaking one should say here that the hypothesis that there is no interdependence between the two sets of rankings has not been confirmed. As Woods, Fletcher and Hughes (1986) point out, it is usually "... difficult to interpret [Spearman's rank correlation] as a measure of 'degree of interdependence'..." (p. 174).
5. To calculate their test-retest reliability score, Jacobs et al. (1981) assessed the performance of two groups of students (size not reported) who took writing tests with two different tasks; the authors found a correlation coefficient of .72 between the mean scores for each task, a figure that they characterize as being "... in effect a test-retest reliability coefficient." (p. 73). Without specifying exactly how many readers took part in their experiment, Jacobs et al. (1981) explain that the "... ranges of reliability coefficients over subsamples of sets of readers who read at least 30 papers each were, for two readers, .59 to .96; for three readers, .89 to .94; and for four readers, .92 to .94" (p. 69). The figure cited in the main text is therefore an *average* interrater reliability score.

References

CITO (1984). *Schrijftoetsen Engels Havo/VWO.* Arnhem, Holland: CITO.

Homburg, T. J. (1984). Holistic evaluation of ESL compositions: Can it be validated objectively? *TESOL Quarterly 18*, 87-107.

Jacobs, H. L., Zinkgraf, S.A., Wormuth, D.R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach.* Rowley, Mass.: Newbury House.

Larsen-Freeman, D. (1983). Assessing global language proficiency. In H. W. Seliger and M. Long (Eds.), *Classroom oriented research in 2nd language acquisition* (pp. 286-305). Rowley, Mass.: Newbury House.

Melse, L. (1984). Toetsing en beoordeling van schrijfvaardigheid in de moderne vreemde talen. *Levende Talen*, 353-360.

Melse, L., & Verstralen, H. H. F. M., (1986). *De ontwikkeling van schrijftoetsen voor de moderne vreemde talen.* Specialistisch bulletin nr. 49. Arnhem, Holland: CITO.

Perkins, K. (1980). Using objective measures of attained writing proficiency to discriminate among holistic evaluations. *TESOL Quarterly 14*, 61-70.

Perkins, K. (1983). On the use of composition scoring techniques, objective measures, and objective tests to evaluate ESL writing ability. *TESOL Quarterly 17*, 651-671.

Woods, A., Fletcher, P., & Hughes, A. (1986). *Statistics in language studies.* Cambridge: Cambridge University Press.

Appendix A
Characterization of Error Types (based on CITO, 1984)

Incompleteness: All of the letters consist of a number of sub-assignments specifying, in Dutch, the contents of the letter. If parts of the sub-assignment have been omitted or carried out improperly, students lose points.

Absolutely Incomplete: Students lose points for not having carried out a sub-assignment at all.

Wordgroup Error: This category refers to wordgroups that contain a complex of errors such that it is difficult to establish exactly how many errors have been made. CITO (1984) gives the following example:
I wanted to be that you are writing about yourself. (p. 15)

Primary Errors: This category contains a considerable variety of errors, most of which are self-evident. Therefore only the less obvious ones will be discussed here.

Excessive repetition: In the example below, the writer is penalized for using the "it was nice" construction with excessive frequency:

> It was nice to see you. It was also nice to see your parents and your little brother. And it was nice to be in England again. (CITO, 1984, p. 15)

Illogical connection: This is a discourse error and results from inadequate textual cohesion:

> We like English t.v. programmes. Apart from that we often watch them.

Lack of clarity:

> When you arrive at the station, I'll be waiting for you at the door. (CITO, 1984, p. 16)

Stations tend to have many doors — and exits, which is presumably what the original author meant — and by not specifying which one is being referred to, the writer is being insufficiently clear about where the meeting is to take place.

Use of word or expression inappropriate to context: These are errors of register, such as:

> Further to our telephone call. . . . (in an informal letter)
> All the best, greetings from. . . . (in a formal letter)
> (CITO, 1984, p. 16)

Too sudden change of subject: When changes of topic are inadequately sign-posted, they are penalized. "_V_" indicates something is missing.

> I have two pets, a cat and a dog, and I like them very much. They always sleep in my bedroom. _V_ My grandfather always snores when he's sleeping. (CITO, 1984, p. 16)

Secondary Errors:

Spelling errors: self-evident.

Punctuation errors: self-evident.

Secondary preposition error: If a student has used an incorrect preposition, this normally means adding, changing, or removing a word in the correction process, and consequently one would have to call it a primary error. CITO, however, makes an exception here:

They haven't looked for the child very well.
John was sitting at the back of Mary.

CITO (1984) points out that in the letters from which these examples came, it was clear that students had intended to writer "after" and "behind" respectively. The first error, however, is counted as a primary one as "to look for" and "to look after" have completely different meanings. This affects comprehension much more than the second error, which does not have a meaning of its own that could confuse readers.

Appendix B
Letter Marked According to CITO Method

Brugstraat 2
Amsterdam
The Netherlands

January 25, 1987

Dear David,

How have you been since my last letter? I'm writing again
because I'm wondering if you have got my pictures 3 months ago.
I really want to know if you liked them. What did you think of
5 our house and my family? I am not very good at taking pictures, but
I hope you got an impression of what our place looks _V_.
Maybe it would be a good idea for me to pay a visit ~~at~~ the
O photography club in our neighbourhood to learn how to take better
pictures, that would also be nice for you.
10 Last week there has been took pictures of our class. Can you
see me standing in the middle row? The tall boy next to me is Daan
Jansen, my best friend.
I hope you will send me a letter soon. Could you please send
me some pictures too? I'm looking forward to hearing from you.
15 Yours sincerely,

Richard

kind of error	**correction**
line 3 grammatical error	Change to "got"
line 6 word needs to be added	Add "like"
line 7 secondary preposition	Change to "to"
line 9 punctuation	Change to full stop + capital letter or to semi-colon.
line 10 wordgroup	It is difficult to see exactly how many corrections would be required here to get, for example, the more acceptable "some pictures were taken."
line 15 use of word or expression inappropriate to context	This being an informal letter, one has to change the expression used to, for example, "best wishes."
paragraph 3	This paragraph is incomplete because the sub-assignment here specified that the student should explain how he had heard of the photography club. The writer of the letter above gives no such explanation, and therefore the paragraph has to be treated as incomplete.

Point value of errors: –7-3/4 Final score: 32-1/4
(letter and some of the comments adapted from CITO, 1984, p. 27)

JALT Journal, Vol. 10, Nos. 1 & 2

TO SEE OUR TEXT AS OTHERS SEE IT: TOWARD A SOCIAL SENSE OF COHERENCE

Daniel Horowitz

Abstract

This paper critically examines the relevant literature on the cognitive processes underlying reading and writing and claims that the essential common point is "interpretation of written text." The concept of interpretation is then examined and it is argued that an understanding of the interpretive standards of the community one is writing for is a necessary prerequisite of the ability to produce coherent, reader-based prose. Some evidence suggests that both L1 and L2 writers gain this understanding through extensive reading, but it is argued that intensive reading and analysis of specific genres can also make an important contribution to writing ability.

Introduction

The phrase "social sense of coherence" expresses the dialectical nature of writing, the tension between the writer and audience. Put simply, it means that writers strike the best bargain they can as they (1) attempt to reconcile what they want to say with what they think their readers are willing to attend to, and (2) attempt to reconcile how they want to say it with the discoursal demands of the genre in which they are working (Swales & Horowitz, 1988). A difficult task, this, and one in which success is inconceivable without a clear understanding of the expectations of the community one is writing for.

How is this understanding acquired? A growing body of research, mainly in L1 but more recently in L2 as well, points toward reading as one key factor. For example, a number of researchers have found a positive correlation between the amount of pleasure reading that

---

Daniel Horowitz teaches in the English Language Program at International Christian University. His main interests are academic reading and writing, testing and computer-assisted language learning. His articles on these subjects have appeared in *TESOL Quarterly, The English for Specific Purposes Journal, Cross Currents,* and *The Language Teacher.*

language learners do and their writing ability (Krashen, 1984 [a review of relevant studies]; Janopolous, 1986). Several writing theorists have gone even further and claimed that extensive reading is a necessary precondition for skilled writing. Krashen (1984), for one, claims that "it is reading that gives the writer the 'feel' for the look and texture of reader-based prose." Flower and Hayes (1980) state it this way: "a well-read person simply has a much larger and richer set of images of what a text can look like" (p. 28).

Implicit in this approach is the idea that reading and writing are closely related. Although the exact nature of that relation is still unclear, recent research (Folman, 1988; Sarig, 1988) suggests that the cognitive processes underlying the two have much in common, and it is the purpose of this paper to speculate on where that commonality may lie so that it may be exploited in the teaching of writing.

Construction of Meaning

The phrase that echoes through many recent descriptions of the cognitive processes of both readers (Beck & Carpenter, 1986) and writers (Emig, 1977; Murray, 1978; Raimes, 1985) is *construction of meaning.* The writer constructs meaning — and in the process is said to discover it — in the act of writing, and the reader re-constructs and re-discovers that meaning by bringing world knowledge to bear on the written symbols left behind by the writer as a "partial record of a discourse enacted by the writer and an imagined interlocutor" (Widdowson, 1986, pp. v-vi).

Although the text itself **physically** links the writer and reader, their shared interpretive abilities enable the text to serve its communicative function. Thus, in order to find what reading and writing have in common, it will not be fruitful to concentrate solely on text itself, for as Widdowson (1986) has stated, "the text, the actual appearance of signs on the page, does not itself contain meaning but provides the occasion for meaning to be achieved in the act of reading" (p. v). Rather, we must examine what is known about how texts and interpreters of texts — writers or readers — come into congruence (or fail to); that is, how a writer infuses a text with potential meaning and how a reader is able to realize that potential.

This somewhat abstract description can be brought into clearer focus by considering how much our ability to write depends in a practical sense on the same abilities which enable us to interpret others' texts. Indeed, for most of us writing is a difficult and even painful process. Few of us know just how our words will look and feel until we see them, and we are sometimes surprised, pleasantly or otherwise, at the reverberations of meaning which thy set off in us. Though the words come out of us, before long they seem to take on a life of their own, and thereafter most of us expend great amounts of energy disciplining them to do our bidding. In that act, as we sense what our own writing means, as we interpret it, recast and re-interpret it, repeating this process until we are satisfied that we know what we want to say and that what we have written means just that, we are engaging the ability to interpret text which we gained by interpreting texts — that is, by reading.

Thus it appears that **interpretation of written text** — our own or others' — is the cognitive process that reading and writing have in common. Such interpretation is possible only when written linguistic symbols and human cognitive structures cohere; in other words, meaning can be constructed only to the degree that coherence exists between text and interpreter. Readers, who cannot change the texts they are presented with, must adapt themselves to those texts as best they can in order to carry on this process. Writers can manipulate text at will but risk alienating their readers if their sense of coherence is too idiosyncratic. Thus, coherence, which is usually construed as either text-based or reader-based (Johns, 1986), is in the present view seen neither as a fixed quality of text nor as a characteristic of a reader's schematic knowledge, but rather as the evolving relation of congruence between these two, as the goal toward which both readers and writers strive in their acts of interpretation. And, to bring the discussion around full circle, it seems clear that one can develop a social (as opposed to idiosyncratic) sense of coherence — see one's text as others will see it —only by becoming skilled in the interpretation of a wide variety of others' texts.

The next section of this paper will attempt to analyze the process of interpretation more closely. The discussion will be cast in such a way that it applies equally to the reading and writing process, but where there are clear differences between the two, these will be noted.

Interpretation

The concept of interpretation as used in the previous discussion rests on the assumption that meanings are "derived" from rather than "asserted" by the text (Keenan, 1978, p. 23). Since interpretation requires the active construction of meaning (as opposed to the passive reception of meaning), this further implies that there are "gaps" between the text itself and the meanings derived from it. The bridging of these gaps is called "inferencing," and interpretation can be seen as the act of using the information gained by inferencing to create a more and more complete picture of the text, or, in the case of writing, to understand what must be added to it or changed in order to create a more complete picture for the intended reader.

If, indeed, the process of interpretation is totally dependent on as potentially unreliable a process as inferencing, and if, as a consequence thereof, different readers interpret a given text in various ways (which we know to be the case), we are faced with the very practical question of where to draw the line between acceptable and unacceptable interpretations. As teachers we face this problem whenever we judge the correctness of our students' interpretations of class readings or the interpretability of student-produced texts. Widdowson (1986) highlights the reading side of this problem when he says that "having rejected the notion that reading is only a matter of discovering meanings which are linguistically encoded in the text, it will not do to go to the other extreme and claim that reading is a matter of unconstrained interpretation subject only to the whims of the individual" (p. vi).

This solution is not entirely satisfactory because it assumes that the "correct" interpretation of a text is the one that conforms most closely to the writer's original intention. Because it makes this assumption, it fails to shed light on the interpretation — or failure of interpretation — of those texts which are ambiguous, either intentionally (certain types of literature or diplomatic discourse) or through poor construction (those written by unskilled native or non-native writers, text written by children, etc.). It also fails to account for those odd but not rare cases when a reader reaches a correct conclusion (perhaps as indicated by answering a test question correctly) based on a highly idiosyncratic understanding of a text.

The crux of the problem is this: Given two interpretations of a text, need we appeal to a writer's intentions to decide which is correct? How do we deal with the situation in which we intuitively feel that there is more than one reasonable interpretation? And how do we justify "overruling" our student-writers who claim that their text means one thing when we are quite sure it means another. . . or nothing at all?

A possible answer to these questions comes from Perelman and Olbrechts-Tyteca (1969), who discuss a similar problem in argumentation. They distinguish between two types of arguments: mathematical ones, in which the conclusion follows deductively from the propositions and of which it can be said that if two people disagree about the validity of the conclusion, one of those people is wrong; and those arguments which one more commonly encounters in real life, where intelligent, well-informed people come to different conclusions based on the same evidence.

Adapting this idea, we can say that there is a continuum of inference and interpretation, at one end of which are those inferences (which generally correspond to lower levels of meaning) on which there will be near unanimity of agreement among some community of readers for whom a given text was intended, and at the other end of which are those inferences (which generally correspond to higher levels of meaning) for which disagreement would be the norm. An example of the lower extreme would be the meaning of a pronoun reference in an academic text; an example of the latter, an interpretation of the symbolism in some difficult literary work.

This continuum is closely related to the notion of inter-rator reliability, a measure of the level of agreement among a group of readers chosen to interpret a text or series of texts. This level varies with the type of text and interpretation required, but it is rarely 1.0, perfect agreement. Likewise, every inference or act of interpretation falls somewhere on the **continuum of reader intersubjectivity**. In the case of reading, this notion allows us to make judgments about the relative or intersubjective correctness of inferences and interpretations without having to make any assumptions about what went on at some time in the past in the mind of a writer. In the case of writing, it clarifies what is meant by a "social" sense of coherence: the internalization by a writer of the interpretive standards of the community he or she is writing for.

Up to this point, this picture of interpretation has been presented: The interpreter constructs meaning by drawing inferences from written symbols (and, in the case of a writer, additionally by changing those symbols to suit his or her evolving intention). In the case of a skilled and qualified reader dealing with a well-constructed text in a familiar genre, the great majority of those inferences will be in agreement with those made by other, similarly qualified members of the reading community, but there may be some which are not; the proportion of agreement to disagreement would tend to be reversed as the reader's skill or familiarity with the genre or subject decreases, or as the text becomes more ambiguous. If, based on these inferences, the reader is able to construct a personally satisfying picture of the text as a whole, we can say that interpretation has taken place, and though we can not say whether that picture corresponds to the writer's original intention, we can involve the notion of intersubjectivity in order to judge in a relativistic way whether that interpretation is more or less correct. Likewise, in the case of a competent writer attempting to create an unambiguous text, the creative and interpretive processes go hand-in-hand until (ideally) the text both matches the writer's intention and conforms to the expectations of his or her community of readers.

In speaking of the interpretive standards of the community a writer is writing for, one is speaking at least in part of the knowledge which that community brings to the interpretive process. The role of knowledge in this process is the subject of the next section of this paper, in which the interpretive process is examined in greater detail using some insights gained from the study of artificial intelligence.

The Role of Knowledge

We begin with the idea that interpretation is a problem-solving process. By "solution" in this case we mean finding satisfactory matches between the highly patterned data coming into the brain from the eye and the patterns of knowledge already stored in long-term memory. In other words, the text input sets off searches for pattern matches through the huge data base of knowledge of the brain. It is clear that in order for these searches to take place in real time, they cannot be random; indeed, "the key to intelligent problem solving lies in reducing the random search for solutions" (Lenat, 1984, p. 152).

What makes these searches possible is, first, the non-random arrangement of knowledge in the brain, and second, the power of the searching "program." The term most commonly used to describe the organized webs of associations which characterize the storage of knowledge in long-term memory is "schemata" (Bartlett, 1932; Neisser, 1976). Text input activates the search for the most appropriate schemata through a powerful searching program, the seat of meta-linguistic knowledge, which services the data base in at least the following ways:

1. It receives raw input from the input device (the eyes) and converts it into a usable form, perhaps by extracting elements from the raw input which are similar to the descriptors used to search a computer data base.
2. It searches the present data base to find the schematic configuration(s) which fit(s) the converted input most closely.
3. It sets levels of probability for deciding if input fits a given schematic configuration.
4. It keeps track of possible "candidate" schemata if the input is still insufficient to make a determination of the most suitable one(s).
5. It directs the search for new raw input to confirm or disconfirm the appropriacy of "candidate" schemata.
6. It decides the degree of fit of input to schemata.
7. It decides, in light of #6, whether or not to modify the existing data base and whether or not to modify itself (in the sense of forming new meta-linguistic rules, creating new searching strategies, setting new probabilities, etc.).
8. It acts recursively, combining the results of lower level searches with more raw input from the input device in order to perform new, higher level searches.
9. It monitors when the process has broken down due to lack of sufficient input, lack of a powerful enough searching strategy, or lack of a sufficiently large data base.

In the act of interpretation, then, text input is converted into instructions for searching and potentially modifying a data base. In the case of reading, to achieve global comprehension of the text, this process spirals upward recursively in the search for higher and higher level schemata which fit longer and longer stretches of text. It also spirals downward, directing the input mechanism to search for a limited number of possible completions of partially established patterns — this is what is meant when we say that schemata create expectations. In the process, schemata are modified as new information or new relations among existing information are added to the data base. In the case of writing, the same process takes place with the obvious difference that the text itself is fluid. From descriptions of the writing process, however, it is clear that as crucial a difference as this is, it does not overshadow the fact that a writer making decisions about the direction of an evolving text does substantially the same things that a reader does in attempting to comprehend one: in both cases meaning is discovered as it is constructed, and in both cases, the ultimate goal is to achieve coherence between an interpreter and a text.

Pedagogical Implications

Swales (1983) provides an example of how "reading-into-writing" works in the classroom. He analyzed a sample of 48 introductions to research articles from various fields, from which emerged a common pattern of four discourse "moves": establishing the field, summarizing previous research, preparing for the present research, and introducing the present research. He further elaborated the model by showing alternative ways the work of each move can be accomplished and by listing some of the common lexical and grammatical patterns found in each move.

He then produced classroom materials designed to familiarize students with these patterns of discourse. These exercises included "colour-coding. . . (in various degrees of detail) the structure of Article Introductions," "jumbled introductions," some with the moves and others with the sentences out of order, and specific language work on some of the structures commonly found in each move. These exercises led to the actual writing of introductions, first as "cloze exercises on whole introductions," then "inserting references into introductory arguments," and finally writing introductions "based on library research cards plus title or abstract" (pp. 197-198).

Swales' work clearly illustrates how reading and writing can — and indeed must — be taught together if students are to internalize the standards of the "rhetorical community" (Purves, 1986, p. 39) they are writing for. Not all such work need be as elegant as Swales' text analysis, however. Teachers can do a great deal of good simply by helping students become aware of the need to "see their texts as others see them" and by introducing the "reading-in-writing" paradigm into their classrooms.

Many currently popular techniques in the teaching of writing are compatible with the first goal. Emphasis on revision and on peer editing — a "real" set of outside eyes — are surely in line with it, though teachers should not forget the important role they play as readers with much more highly developed writing schemata than their students. Encouraging students to let their writing "sit" for a few days before revising it also heightens their ability to see their text as others see it.

At an even more basic level, teachers can work within their institutions to promote the integration of reading and writing. Unfortunately, there still seems to be a strong tendency to think of reading and writing as separate skills and to teach them in separate classes. Though there may be good reason to devote a separate class to the teaching of reading strategies, there is no justification for teaching writing without a strong "reading for writing" component. Readings should serve two main functions in a writing class, as sources of facts and ideas (Horowitz, 1986) and as models of the type of writing students will be expected to do. Unfortunately, it is the rare writing textbook or even rhetorical reader which presents students with models of writing **based on the types of texts they will eventually be required to produce** (other than personal essays) along with exercises designed specifically to build up the schematic knowledge of genre which leads to critical reading and, in turn, to successful writing. This may simply be a reflection of the economic realities of textbook publishing, but whatever the reason, it is up to teachers to make up for this lack by finding out what types of texts their students will have to produce and by designing exercises to help them become informed, critical readers of their own writing.

Conclusion

By careful analysis of the tasks our students will eventually face and the texts they will have to read and produce, by judicious selection and imaginative creation of classroom materials, and by acceptance of one's role as the students' window on the rhetorical community they are soon to join, teachers of reading and writing can guide their students to the goal of all reading and writing instruction — a social sense of coherence. There is no more we can do, and our students deserve no less.

Acknowledgments

I want to thank Ann Johns for her helpful comments on an earlier draft of this paper, and Eriko Machi for her help in the preparation of the abstract.

References

Bartlett, F. (1932). *Remembering.* Cambridge: Cambridge University Press.

Beck, I. L., & Carpenter, P. A. (1986). Cognitive approaches to understanding reading. *American Psychologist, 41,* 1098-1105.

Emig, J. (1977). Writing as a mode of learning. *College Composition and Communication, 28,* 122-128.

Flower, L., & Hayes, J. (1980). The cognition of discovery: Defining a rhetorical problem. *College Composition and Communication, 31,* 21-32.

Folman, S. (1988). Towards an EFL reading-writing model of academic learning. Paper presented at the 22nd Annual TESOL Convention, Chicago, IL.

Horowitz, D. (1986). What professors actually require: Academic tasks for the ESL classroom. *TESOL Quarterly, 20,* 445-462.

Janopoulos, M. (1986). The relationship of pleasure reading and second language writing proficiency. *TESOL Quarterly, 20,* 763-768.

Johns, A. M. (1986). Coherence and academic writing: Some definitions and suggestions for teaching. *TESOL Quarterly, 20,* 247-265.

Keenan, J. M. (1978). Psychological issues concerning implication: Comments on "Psychology of pragmatic implication: Information processing between the lines" by Harris and Monaco. *Journal of Experimental Psychology: General, 107,* 23-27.

Krashen, S. D. (1984). *Writing.* Oxford: Pergamon.

Lenat, D. B. (1984). Computer software of intelligent systems. *Scientific American, 251,* 152-160.

Murray, D. M. (1978). Internal revision: A process of discovery. In C. R. Cooper & L. Odell (eds.), *Research on composing: Points of departure* (pp. 85-103). Urbana, IL: National Council of Teachers of English.

Neisser, U. (1976). *Cognition and reality.* San Francisco: W. H. Freeman.

Perelman, C., & Olbrechts-Tyteca, L. (1969). *The new rhetoric: A treatise on argumentation* (J. Wilkinson & P. Weaver, Trans.). Notre Dame and London: University of Notre Dame Press.

Purves, A. C. (1986). Rhetorical communities, the international student, and basic writing. *Journal of Basic Writing, 5,* 38-51.

Raimes, A. (1985). What unskilled ESL students do as they write: A classroom study of composing. *TESOL Quarterly, 19,* 229-258.

Sarig, G. (1988). Composing a study-summary: A reading-writing encounter. Paper presented at the 22nd Annual TESOL Convention, Chicago, IL.

Swales, J. (1983). Developing materials for writing scholarly introductions. In R. R. Jordan (ed.), *Case studies in ELT* (pp. 188-200). London: Collins.

Swales, J., & Horowitz, D. (1988). A genre-based approach to ESL materials. Paper presented at the 22nd Annual TESOL Convention, Chicago, IL.

Widdowson, H. G. (1986). Foreword. In S. Salimbene, *Interactive reading* (pp. v-vi). Rowley, MA: Newbury House.

POINT TO POINT

PARAPHRASE, CONTEMPORARY POETRY, AND LITERARY SYLLABUSES IN JAPAN.

Richard Cauldwell

Allan Hirvela's article (*JALT Journal* 9.2) on the integration of language and literature teaching raises interesting questions of which I should like to address three:

—Can poetry be taught using paraphrases?
—In what way is poetry suitable for literature courses in Japan?
—What should teachers and students do with poems?

1. Can Poetry Be Taught Using Paraphrases?

Alan Hirvela (1988) implicitly excludes poetry from the type of treatment he suggests for non-native speakers studying literature:

> Poetry, by its very nature, cannot really be simplified within the same literary form or genre. To be sure, poems are simplified through paraphrasing in prose form; however, comparisons of these vastly different representations of the text would be difficult. . . . (p. 140)

I will argue that teaching students to **create** first a spoken, then a written paraphrase of a poem, can be an effective basis for a literature course. Any commentary on a poem, by a literary critic or a high-school student, involves elements of paraphrase, and learning how to paraphrase is an important literary skill.

My goal, in my English literature courses at a national university in Japan, is not to compare two already prepared versions of the same text; rather it is to help the students to analyze a poem and produce a paraphrase of it. This paraphrase will consist of talk and writing about the structure, patterns, and meanings of the poem. A second aim is to teach the students a strategy for paraphrasing which they can apply to other poems.

2. How Suitable Is Poetry?

There seems to be a general assumption that poetry is "difficult" for native, and therefore "too difficult" for non-native speakers. However, it is my experience that many poems by some contemporary and recent poets make ideal texts for literary courses at university level in Japan.

Many of them are short (under thirty lines) with relatively easy vocabulary, addressing themes common to the literatures of many cultures (love, growth, loss). They are in regular stanzas, have rhyme schemes, and are not dense with metaphor (e.g. Seamus Heaney's "Scaffolding," where a building under construction is a metaphor for a developing relationship). They have a readily visualizable physical or psycho-social setting (e.g. Tony Harrison's "Book Ends," in which a bereaved husband and son sit in front of a fire). Finally they often are written in everyday speech (e.g. the questions in Causley's "Whatever Happened to Lulu"; or Auden's colloquialisms in "The More Loving One").

Poems with these characteristics can be analyzed and paraphrased by the students in a ninety-minute session. Vocabulary problems can be overcome within half an hour; then literary qualities and overall meaning may be addressed. Each lesson can thus be free-standing and independent, yet be one of a series which contributes to overall course objectives. This flexibility is vital in situations such as mine where special lectures and other requirements made of students mean that their attendance over a three-month, twelve-session course is unpredictable.

One argument against using such recent poems is that they are not yet part of the literary canon: Posterity has not made its judgment on this body of work; knowing about this small corpus of contemporary British poetry will be of little use to the students, even those with literature majors. Against this one has to weigh the advantages of providing students with a more direct experience of primary, if not canonical, texts.

Poems that have proved both popular and useful for my students include Derek Mahon's "Leaves" and "The Snow Party" (enjoyed for its mention of Japan); Tony Harrison's "Long Distance" and "Book Ends". These and other suitable poems can be found in Morrison and Motion (1982). A second anthology is Maley and Moulding (1985), which contains poems by Charles Causley, Seamus Heaney, Eleanor Farjeon, and others. Journals and publications such as the *Times Literary Supplement* are other good sources.

3. What Should Teachers and Students Do With Such Texts?

Even with appropriate poems, the students will gain little if the teacher then overwhelms them with inappropriate discourse such as long difficult lectures or written critiques of the poems. Because students must understand poems before they can produce a paraphrase, the discourse with which poems are approached has to be more accessible than the poems themselves.

The most accessible talk and writing is that which the students produce themselves. In order to get students to produce this language, I favor having students in horseshoe-shaped groups of five (I have done this with up to forty people) with the open end of the group towards me. Thus we can switch from lecture-mode to intra-group discussion very quickly. I give short lectures of about three minutes, or, if longer, broken up into two or three chunks of about three minutes in length. I then ask each group, under the direction of a group leader, to check if they understood the content of the lecture, and come back to me with questions if they feel they do not. I usually direct them to speak in English: occasionally I offer them a choice of either Japanese or English, if the target concept is beyond the English ability of some. The lecture might consist of an introduction to the poet or to the literary qualities being addressed in that session (rhyme, rhythm, metre).

The groups then start to analyze the poem, and if they have not done vocabulary work on the poem beforehand, they check words they feel they do not understand. Once students understand the poem word-for-word, we address its structure, its patterns, and overall meaning. I give students tasks on the poem. In the early stages of a course the task I give is very specific: "count the number of words which have prominent /s/ sounds" ("Book Ends"); "How many syllables in each line?" ("Scaffolding"). Later on, as students become more aware of the variety of possible literary patterns, the task I give is more general "Look for patterns"; "How do these patterns relate to the meanings?"; "Who is speaking to whom? What evidence does the poem provide for us to identify the persona of the poem or the person s/he is addressing?"; "Why this title?" ("Long Distance"; "Book Ends").

The sessions are driven by a four-stage cycle of:

(1) Lecturer setting the task
(2) Students performing the task
(3) Students reporting back
(4) Lecturer providing language.

Each cycle varies in length from two to a maximum of fifteen minutes. After each task is performed, the group leader reports on the group's findings (the leadership is passed to a different student for each cycle).

During the performance of the task and reporting back, the need for providing the students with the language necessary to talk about literature becomes apparent. Usually they are trying to express meanings for which they do not possess the appropriate language. This includes not only technical terms (iambs, assonance, alliteration, pararhyme, etc.) but also how to: refer to specific parts of the poem ("in the second line of the third stanza"); enumerate the properties of the poem ("this poem consists of 7 four-line stanzas rhyming ABAB"); comment on patterns and deviations from patterns ("all the lines of this poem contain 10 syllables, except for the last which contains 9"); and relate these structural statements to statements about the meaning of the poem ("the predominance of /s/ sounds in the poem suggests..."). In addition they need to learn strategies such as hedging: "one could even argue that... but..."; and could even learn how to use a metaphor from the poem to talk about the poem (a common strategy of critics; for an example see Thom Gunn, 1988).

4. Conclusion

To use a poem to generate "interesting discussion" is, arguably, appropriate for a language course. But literature courses demand that students learn things literary. Part of the purpose of any literary class must be to learn acceptable ways of discussing literature. I argue that teaching students how to paraphrase a poem, both in speech and writing, is indeed a worthy basic aim for any literature course.

It is not enough to expose students to literary works, to give students experience of poems and invite free discussion; for one thing, this makes testing very difficult. Students have to learn how to speak and write about poetry. In short, they have to learn not only how to experience literary discourse in the form of poems, they have to learn how to produce literary discourse in the form of descriptive and evaluative criticism and appreciation. Exposing students to literary works without showing how to produce literary discourse is to fall into the input-only heresy.

A literary syllabus should therefore include three types of discourse, all three of which have to be adjusted to the students' level: the texts themselves (in this case I have argued the case for contemporary poetry), the spoken discourse which is necessary for the analysis and commentary on the poems in class, and the written discourse necessary for assessment.

References

Gunn, T. (1988). February 19-25. Inventing the completely new poem. (A review of volume 1 of the collected poems of William Carlos Williams). *Times Literary Supplement*, 179-180.

Hirvela, A. (1988). Integrating simplified and original texts. *JALT Journal 9*, 131-151.

Maley, A., & Moulding, S. (1985). *Poem into Poem.* Cambridge: Cambridge University Press.

Morrison, B., & Motion, A. (1982). *The Penguin Book of Contemporary British Poetry.* Harmondsworth: Penguin.

NEW FROM PHR

DISCOURSE ACROSS CULTURES

Strategies in World Englishes

EDITOR: LARRY E. SMITH,
East-West Centre, Honolulu, Hawaii

This pioneering volume consists of thirteen specially commissioned papers on theoretical, applied, and pedagogical aspects of the strategies of discourse in selected world varieties of English. It presents a new perspective and fresh insights for understanding, teaching, and research related to English in international contexts.

The book is designed for those interested in theoretical and applied linguistics, language teaching, and literatures written in English. The five sections of the book are:

Part 1. Literature, Conventions and Explanation.
Part 2. Discourse Expectations and Cross-Cultural Communication.
Part 3. Text, Context and Culture.
Part 4. The Bilinguals' Creativity and Contact Literatures in English.
Part 5. Culture and the Language Classroom.

In his foreword, Sir Ralph Quirk observes that this volume *"deserves special attention, because it focuses on two unique functions of English. It not only discusses the role of English as an international language, but it also presents the international uses of English in various literatures around the world."*

13-215005-0

NEW TITLES IN METHODOLOGY FOR 1988 FROM PRENTICE HALL

Problems into Action
Troubleshooting in the English Language Classroom
C. Candlin, G. Gerngross & H. Puchta

A vade mecum for the EFL teacher and teacher trainer, with its easily accessible and readable focus on the real problems of daily teaching in the light of current principles and practice.
Identifying specific problem areas for classroom language teachers and students, the authors provide guidance towards explicit techniques through carefully structured stages of exploration, exemplification and discussion.
13-716317-7

Language Teachers at Work
ANTHONY J. PECK
University of York

In a clear way, this book shows how a number of classroom teachers have taught specific lessons in French, German and English as foreign languages.
The descriptions and analyses will allow the reader to focus on systematic classroom methodology through a comparison of the case-studies presented.
The emphasis of the classes on • how teachers present new texts • how they conduct oral work • how they develop the learner's free expression • how they teach grammar and • how they teach listening comprehension.
13-523190-9

Language Planning and Language Teaching
C. KENNEDY
University of Birmingham

So far the links between Language Planning and what happens in ELT classrooms around the world have not been made explicit.
The articles in this collection have been carefully chosen to illustrate the connections between Language Planning and Language Teaching with illustrations drawn from a variety of ELT contexts in Africa and Asia. Together the articles raise a number of important issues –

- Can second language acquisition be planned?
- What is the connection between the role of English and ELT materials?
- What should the language of instruction be?
- When should English be introduced on the curriculum?
- What variety of English should be taught?

13-523184-1

For further information, contact Harry T. Jennings.

PHR
PRENTICE HALL REGENTS

Prentice Hall Regents of Japan
Jochi Kojimachi Bldg., 3F, 6-1-25 Kojimachi, Chiyoda-ku, Tokyo 102, JAPAN
東京都千代田区麹町6-1-25 上智麹町ビル3階
Telephone (03)238-1050 Fax (03)237-1460 Telex 650-295-8390
(For telexes from outside U.S. precede with U.S. Access Code 23 from most countries)

BOOK REVIEW

THE CLASSROOM AND THE LANGUAGE LEARNER. Leo van Lier. Harlow: Longman, 1988. 226 pp. ¥2,340.

Teachers are always on the lookout for new ideas that will work in their particular classrooms. For many hard-pressed teachers with very little time to keep up to date with the constant flow of developments in the field, the exchange of ideas is largely anecdotal, occurring in conversations with colleagues at odd moments in the staff room. These discussions, which are typically concerned with "what works" in a lesson or part of a lesson are not to be disparaged. They constitute one of the few means of professional development available to teachers once their training has finished.

This informal sharing of ideas is, arguably, classroom research, in that conclusions about teaching and learning are drawn from happenings observed in class time. But the conclusions arrived at in such discussions are vulnerable to two different kinds of objection.

The first objection concerns the issue of scope. Teachers understandably prefer ready-to-use techniques and materials that contribute to an enjoyable and purposeful *lesson* and are concerned rather less with their impact on the long-term aims of a *course*. The criteria by which we gauge the effectiveness of new ideas (and old) need to be widened to integrate short-term and long-term objectives.

The second kind of objection stems from the familiar distrust of personal intuition as a means of describing what is happening in the classroom. What one teacher or student sees as success, after all, is open to different interpretations by others.

The call to replace suspect intuitions with more valid means of gathering information is, of course, a call to make descriptions of what goes on in a classroom more *objective*. Leo van Lier's central purpose in *The Classroom and the Language Learner* is to establish ways in which this objectivity can be achieved. Van Lier hopes that success in this enterprise will lead to better classroom practice; and as a corollary, though he never states it explicitly in the book, he hopes that objectivity will enable teachers more clearly to relate the objectives of individual activities to lessons, and to relate individual lessons to a course.

Principled investigation of what actually goes on in classrooms is a relatively new discipline lacking an agreed method of operation. Van

Lier therefore feels justified in devoting the first 89 pages of the book to an analysis of how classroom research differs from research in other fields; how it should be carried out; and what phenomena should be investigated. This is the clearest and most readily useful section of the book, and it allows teachers to begin planning relevant research projects for the contexts in which they work.

Van Lier is very clear about the limitations of the results of such research. Since our understanding of the learning process in and out of the classroom is restricted, it is inappropriate to expect classroom research to produce the kind of causal statements (such as technique "A" will produce effect "B") common in physical sciences. What we can expect is that classroom research will provide not a *complete* understanding but a *better* understanding of what goes on.

Once van Lier has set the parameters within which classroom research can operate, he addresses the issue of how it is best carried out. He adopts the view that classroom research is a help to understanding language learning rather than a means of generating universal theories. He suggests that any piece of research should begin with transcription — a kind of insurance policy against overlooking important phenomena. The task of recording and transcribing the language which teachers and students use is time-consuming, but crucial if classroom research is to make any valid claims to objectivity. The data on which one makes any kind of inference must, after all, be made available for scrutiny.

I found the explanation of how transcription can be facilitated with the use of techniques borrowed from studies in discourse analysis somewhat difficult to follow. Van Lier does provide a helpful appendix on transcribing, but the text itself often seemed overwrought with detail, clouding rather than clarifying the issue.

What of the subjects van Lier considers amenable to the kind of classroom research he has in mind? He begins by making the uncontentious point that teachers and researchers are very likely to highlight some features of the classroom rather than others; and what they choose to focus on will depend on their particular beliefs and hunches about how language is learnt. But here again van Lier diverges from his argument to discuss, albeit briefly, rival theories of language acquisition. When he does return to the point, clarity is partially restored, but still there is no principled alternative to a "let's-describe-anything" approach. The researcher, it appears, in the process of transcribing, will be able gradually to impose order on the data, so that understanding

> . . .emerges from the data rather than from what one already patently knows, and the transcription and analysis process is an aid in developing this emergent order. (p. 81)

Most of the rest of the book is divided into three chapters which, as Professor Candlin, in a preface I otherwise found arcane, helpfully points out, deals with the central theme of the book. That is how, and

> to what point, and to what extent such connections between traces in the described text [the transcription] and these social forces [of the classroom] can be made. (p. x)

To do this, van Lier applies his method to three issues: *interaction, participation* and *repair* (the ways teachers and students deal with error). Given the author's previous insistence on allowing issues to emerge *from* data rather than *before* it, it is interesting to see how these issues arise as well as how his views on classroom research work out in practice.

The first point is resolved immediately. While the initial third of the book suggests that the researcher go into the classroom "cold" and note phenomena as they emerge in transcription. But this is simply not all that he does. Van Lier precedes what might be called the "classroom phase" of research by investigating areas of general interest to all teachers, which in turn generate a range of questions that may or may not be answered when data is examined. In the chapter on interaction, for example, the questions addressed are:

1. How can interaction be exploited to promote language development?
2. If language learning occurs through meaningful interaction, what kind of interaction can be called "meaningful"?
3. How does social interaction in the classroom relate to the cognitive behavior of students?
4. How do students show that they are participating in class?

Van Lier gives a clear account of the different kinds of turn taking that can happen in class so that the researcher or teacher can compare the performance of students with the interaction options taken. It is in this way that van Lier combines the linguistic discipline of discourse analysis with social psychology to produce the basis for sound classroom research. It is at this point in the book that van Lier's insistence that classroom research is a branch of ethnography (literally the study of culture) that must cast its empirical net as widely as possible, is finally made clear.

Both the author and the general editor of this Longman series, Professor Candlin, claim their audience to be teachers as well as researchers, but one can't help feeling the latter group has received more attention than the former. Topics throughout the book are given an historical treatment as explanations emerge from van Lier's discussion of the literature. I would have preferred a more direct account of what constitutes classroom research, with the contribution of other scholars dealt with parenthetically.

On the other hand, the book should be required reading for those teachers and teacher trainers who are interested in getting more reliable — objective — information from the classroom. It also signposts ways of maximizing the kind of interactions in the staff room I mentioned earlier. By extension too, of course, its relevance to teacher training is clear: the more confidence we have in what works in the classroom, the more efficient teacher training can be.

Reviewed by Tom Hinton
Cambridge English School
The British Council, Tokyo

CROSS CURRENTS

AN INTERNATIONAL JOURNAL OF LANGUAGE TEACHING AND CROSS-CULTURAL COMMUNICATION

Cross Currents pursues both theoretical and practical issues central to ESL/EFL instructors. Within each issue, you'll discover articles about English language teaching theory, balanced by ideas for daily lesson planning, all of which is put into international perspective by cross-cultural themes.

Look at selections from recent issues:

- Anderson: "Creative Strategies for Large Class Management"
- Gardner: "Cultural Differences in Non-Verbal Communication: Practical Activities for the ESL/EFL Classroom"
- Scovel: "The Power of English Prepositions"
- Wajnryb: "CLL Revisited: Addressing Socio-Cultural Factors in Language Learning"
- Wongbiasag: "Learner Problems in Self-Access Learning: A Case Study"
- Paulet: "Widening Our Horizons Through Creative Counselling"
- Bright Ideas
- Book Reviews

THE LANGUAGE INSTITUTE OF JAPAN

ISSN 0289-1239

Cross Currents: Written by teachers for teachers

SUBSCRIPTIONS	Single issue	1 year	2 years
Inside Japan*	¥1,250	¥2,500	¥5,000
Outside Japan—Individual Subscription	$8.00	$15.00	$25.00
Outside Japan—Institutional Rate		$19.50	$35.00

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

Cross Currents is published twice a year by the Language Institute of Japan.

- ***Japan Subscriptions:***
  School Book Service
  FI Building 4F
  1-26-5 Takadanobaba
  Shinjuku-ku, Tokyo 160
  03-200-4531
  (By postal transfer to Tokyo postal account No. 9-86192, or cash delivery–Genkin Kakitome.)
- Current Issues are also available in major Japanese bookstores through YOHAN in Japan.
- Back Issues are available from Cross Currents.
- ***Outside Japan:***
  Cross Currents
  Language Institute
  of Japan
  4-14-1 Shiroyama
  Odawara, Kanagawa 250
  Japan
  (Make checks payable to *Cross Currents* (LIOJ) drawn on a U.S. bank.)

INFORMATION FOR CONTRIBUTORS

EDITORIAL POLICY

The *JALT Journal* welcomes submission of articles which contribute to the field of foreign language teaching and learning in both the theoretical and practical domains, especially in the following areas:

1) curriculum, methods and techniques
2) classroom observation
3) teacher education and training
4) cross-cultural studies
5) language learning and language acquisition
6) overviews of research and practice in related fields

In particular, the editors encourage submissions which examine issues of research and practice within the Japanese context, although articles of interest to an international audience are always appreciated. The editors also invite the contribution of short articles, as well as commentary on material that has appeared in previous issues of the *Journal.*

The *Journal* would prefer all articles that are submitted to be written with a general, language education audience in mind. So if you have used certain statistical techniques, or generally unfamiliar terms, please devote more space to explaining the techniques and terms than you would need to when writing for a more technical publication.

INFORMATION FOR AUTHORS

Manuscripts should be no longer than 30 pages, typed on A4 paper and double-spaced. The *Journal* conforms to the style of the *The Publication Manual of the American Psychological Association* (third edition) which can be obtained from the Order Department, American Psychological Association, 1200 Seventeenth Street, N.W., Washington, D.C. 20036. The following may be considered a general guide. Reference citations should be placed in the body of the text in parentheses with the author's last name, date of the work cited, and page numbers. Footnotes for substantive information should be minimized. Authors are responsible for complete and correct information in the reference list and block quotations.

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

Submit three copies of the manuscript, an abstract of less than 200 words, a running head title of about 5 words, and a biographical sketch of not more than 50 words. Wherever possible, the editors would be pleased to receive Japanese translations of the abstracts.

Papers submitted to the *Journal* should not have been published previously, neither should they be under consideration for publication elsewhere. We regret that manuscripts cannot be returned.

Apart from book reviews, all manuscripts and inquiries should be directed to Richard Cauldwell, *JALT Journal* c/o JALT Central Office, Kyoto English Center, Sumitomo Seimei Bldg., Shijo-Karasuma Nishi-iru, Shimogyo-ku, Kyoto 600. (075) 221-2376.

BOOK REVIEWS

If you wish to submit a book review, please contact the Book Review Editor at the address given below before writing the review. A list of publications which have been sent to JALT for review is published each month in *The Language Teacher*. Rita Silver, book review editor, Osaka Jogakuin Junior College, 2-26-54 Tamatsukuri, Higashi-ku, Osaka 540.

ADVERTISING

Inquiries about advertising should be directed to John E. Burton, JALT Publications Business Manager, Suzuki Aparto 4, Nakakaigan 4-3-14, Chigasaki-shi, Kanagawa-ken 253.

JALT JOURNAL　第10巻　第1・2号

昭和63年11月25日　印刷
昭和63年12月1日　発行

発行人　全国語学教育学会(JALT)事務局長
THOMAS N. ROBB

発行所　JALT本部
〒600　京都市下京区寺町仏光寺下ル
ライオンズマンション111号
TEL　(075) 221 - 2376

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

NEW FOR 1988

BBC English courses

THE LOST SECRET is a book and video course intended for users whom, its authors, Robert O'Neill and Martin Shovel, describe as 'false' or 'daring' beginners.

The video consists of an adventure story in 11 parts about a British archaelogist, Dr. Ross Orwell, who is an expert on an extinct South American civilisation, Mepatecs.

With its strong, intriguing story, the video makes use of learners' interest in the unfolding narrative as a means of helping them understand. Although the course begins with very basic vocabulary and structures, the dialogue is constructed in such a way that the use and repetition of key items arise naturally out of the development of the plot.

THE LOST SECRET

Robert O'Neill & Martin Shovel

INTERNATIONAL LEARNING SYSTEMS (JAPAN) LTD.
Matsuoka Central Bldg. 8F, 1-7-1 Nishishinjuku
Shinjuku-ku, Tokyo 160 Tel :03-344-3412
Fax:03-344-3405
Osaka Office: Osaka Ekimae Dai-4 Bldg. 15F.
1-11-4 Umeda, Kita-ku, Osaka 530
Tel:06-347-4733 Fax:06-347-4734

ILS

ADDISON WESLEY READING PROGRAM

READING ON PURPOSE
Fraida Dubin
Elite Olshtain

READING ON PURPOSE

Building Cognitive Skills for Intermediate Learners

Fraida Dubin University of Southern California
Elite Olshtain Tel-Aviv University

READING POWER

BEATRICE S. MIKULECKY Boston University
LINDA JEFFRIES

Reading Faster・Thinking Skills
Reading for Pleasure・Comprehension Skills

READING POWER
Reading Faster・Thinking Skills
Reading for Pleasure・Comprehension Skills
ADDISON-WESLEY PUBLISHING COMPANY

READING BY ALL MEANS

Fraida Dubin／Elite Olshtain
A Developmental Reading Text for Intermediate and Advanced ESL／EFL Students

THREE EASY PIECES

Fraida Dubin／Elite Olshtain
Exciting ESL／EFL Readers Written Especially for the Low Intermediate Level

American Reading Series Life and Culture Edited by Kenji Kitao

AN AMERICAN SAMPLER

AMERICAN PATTERNS

AMERICAN VISTAS

American Patterns

American Vistas

AN AMERICAN SAMPLER

アジソン・ウェスレイ・パブリッシャーズ・ジャパン 〒101 東京都千代田区猿楽町1-2-2 TEL.03-291-4581 FAX03-291-4592
Addison-Wesley Publishers Japan Nichibo Bldg. 1-2-2, Sarugaku-cho, Chiyoda-ku, Tokyo 101