

Japan Association for Language Teaching

Articles

- 5 Readability Formulas For EFL
Jerry Greenfield
- 25 Roles of Output and Feedback for L2 Learners' Noticing
Hideki Sakai
- 55 Using Mnemonics to Facilitate Learning of Japanese Script Characters
Emmanuel Manalo, Satomi Mizutani, and Julie Trafford

Perspectives

- 79 English in Japan: The World Englishes Perspective
Phillip R. Morrow

Reviews

- 101 *Task-Based Language Learning and Teaching* (Rod Ellis) Reviewed by Thomas C. Anderson
- 103 *Methodology in Language Teaching: An Anthology of Current Practice* (J. Richards & W. Renandya [Eds.]) Reviewed by Roger Kenworthy
- 105 *Practical English Language Teaching (International Edition)* (David Nunan [Ed.]) Reviewed by Paul Hullah
- 107 *Teachers' Narrative Inquiry as Professional Development* (Karen E. Johnson & Paula R. Golombek [Eds.]) Reviewed by Michael J. Crawford
- 109 *Controversies in Applied Linguistics* (Barbara Seidlhofer [Ed.]) Reviewed by Scott Petersen
- 111 *Applied Linguistics in Language Education* (Steven McDonough) Reviewed by Scott Bronner
- 113 *The Changing Face of CALL: A Japanese Perspective* (Paul E. Lewis [Ed.]) Reviewed by Justin Charlebois
- 114 *Pragmatics* (George Yule) Reviewed by Donna Tatsuki
- 116 *The Bilingual Family: A Handbook for Parents (Second Edition)* (Edith Harding-Esch & Philip Riley) Reviewed by Jack Massalski
- 118 *The Japanese Mind: Understanding Contemporary Japanese Culture* (Roger J. Davies & Osamu Ikeno [Eds.]) Reviewed by Jonathan M. W. Rankine

JALT Journal

Volume 26 • No. 1

May 2004

Editor

Donna Tatsuki

Kobe City University of Foreign Studies

Associate Editor

Steve Cornwell

Osaka Jogakuin Junior College

Reviews Editor

Kate Allen

*Kanda University of International
Studies*

Japanese-Language Editor

Sayoko Yamashita

Tokyo Medical and Dental University

Editorial Board

David Aline

Kanagawa University

David Beglar

Temple University Japan

William Bradley

Ryukoku University

James Dean Brown

University of Hawai'i, Manoa, USA

Charles Browne

Aoyama Gakuin University

Christine Pearson Casanave

Teachers College, Columbia University, Tokyo

Ann Ediger

Hunter College, City University of New York, USA

Rod Ellis

University of Auckland, New Zealand

Sandra Fotos

Senshu University

Eli Hinkel

Seattle University, USA

Nicholas O. Jungheim

Waseda University

Tim Murphey

Dokkyo University

Yuji Nakamura

Tokyo Keizai University

Mary Goebel Noguchi

Ritsumeikan University

Akito Ozaki

Nagoya University

Aek Phakiti

Maejo University, Thailand

Tim Riney

International Christian University

Carol Rinnert

Hiroshima City University

Thomas Robb

Kyoto Sangyo University

Gordon Robson

Showa Women's University

Tadashi Sakamoto

Nanzan University

David Shea

Keio University

Deryn Verity

Nagoya University of Commerce and Business

Administration

Brad Visgatis

Osaka International University for Women

Minoru Wada

Meikai University

Additional Readers: William Acton, Dwight Atkinson, Eton Churchill, Hideo Hosokawa, Yoshiko Kawamura, Sachiko Kondo, Michael Lessard Clouston, Megumi Kawate-Mierzejewska, Jiro Nishigori, Ethel Ogane, Yoko Okita, Brent Poole, Mark Sawyer, Parrill Stribling, Ken Tamai, Ken Urano, Deryn Verity, Martin Willis

JALT Journal Proofreading: Carolyn Ashizawa, Brian Bresnihan, Megumi Kawate-Mierzejewska, Steve McGuire, Jack Yohay

JALT Publications Board Chair: Amanda O'Brien

Layout: Pukeko Graphics

Cover Design: The Word Works

JALT Journal on the Internet: <http://www.jalt.org/jj>

Website Editor: Anthony Crooks

Contents

May 2004
Volume 26 • No. 1

- 3 In This Issue
- 4 From the Editors

Articles

- 5 Readability Formulas For EFL
Jerry Greenfield
- 25 Roles of Output and Feedback for L2 Learners' Noticing
Hideki Sakai
- 55 Using Mnemonics to Facilitate Learning of Japanese Script Characters
Emmanuel Manalo, Satomi Mizutani, and Julie Trafford

Perspectives

- 79 English in Japan: The World Englishes Perspective
Phillip R. Morrow

Reviews

- 101 *Task-Based Language Learning and Teaching* (Rod Ellis)
Reviewed by Thomas C. Anderson
- 103 *Methodology in Language Teaching: An Anthology of Current Practice* (J. Richards & W. Renandya [Eds.]) Reviewed by Roger Kenworthy
- 105 *Practical English Language Teaching (International Edition)* (David Nunan [Ed.]) Reviewed by Paul Hullah
- 107 *Teachers' Narrative Inquiry as Professional Development* (Karen E. Johnson & Paula R. Golombek [Eds.])
Reviewed by Michael J. Crawford
- 109 *Controversies in Applied Linguistics* (Barbara Seidlhofer [Ed.])
Reviewed by Scott Petersen
- 111 *Applied Linguistics in Language Education* (Steven McDonough)
Reviewed by Scott Bronner
- 113 *The Changing Face of CALL: A Japanese Perspective* (Paul E. Lewis [Ed.])
Reviewed by Justin Charlebois
- 114 *Pragmatics* (George Yule) Reviewed by Donna Tatsuki
- 116 *The Bilingual Family: A Handbook for Parents (Second Edition)* (Edith Harding-Esch & Philip Riley) Reviewed by Jack Massalski
- 118 *The Japanese Mind: Understanding Contemporary Japanese Culture* (Roger J. Davies & Osamu Ikeno [Eds.]) Reviewed by Jonathan M. W. Rankine

JALT Journal Information

- 120 Information for Contributors (English and Japanese)

All materials in this publication are copyright ©2004 by their respective authors.

Japan Association for Language Teaching

A Nonprofit Organization

The Japan Association for Language Teaching (JALT) is a nonprofit professional organization dedicated to the improvement of language teaching and learning in Japan. It provides a forum for the exchange of new ideas and techniques and a means of keeping informed about developments in the rapidly changing field of second and foreign language education. Established in 1976, JALT serves an international membership of more than 2,500 language teachers. There are 40 JALT chapters in Japan, 15 special interest groups (SIGs), and three forming SIGs. JALT is one of the founders of PAC (Pacific Asian Consortium), which is an association of language teachers' organizations in Pacific Asia. PAC holds regional conferences and exchanges information among its member organizations. JALT is the Japan affiliate of International TESOL (Teachers of English to Speakers of Other Languages) and is a branch of IATEFL (International Association of Teachers of English as a Foreign Language).

JALT publishes JALT Journal, a research journal; The Language Teacher, a monthly periodical containing articles, teaching activities, reviews, and announcements about professional concerns; and JALT International Conference Proceedings.

The JALT International Conference on Language Teaching and Learning and Educational Materials Exposition attracts some 1,600 participants annually and offers over 300 papers, workshops, colloquia, and poster sessions. Each JALT chapter holds local meetings and JALT's SIGs provide information on specific concerns. JALT also sponsors special events such as workshops and conferences on specific themes, and awards annual grants for research projects related to language teaching and learning.

Membership is open to those interested in language education and includes enrollment in the nearest chapter, copies of JALT publications, and reduced admission to JALT-sponsored events. JALT members can join as many SIGs as they wish for an annual fee of ¥1,500 per SIG. For information, contact the JALT Central Office or visit the JALT website at <<http://www.jalt.org>>.

JALT National Officers, 2004

President: James Swan
Vice President: Tadashi Ishida
Auditor: Morijiro Shibayama
Director of Treasury: Peter Wanner
Director of Records: Mary Christianson
Director of Programs: Alan Mackenzie
Director of Membership: Hugh Nicoll
Director of Public Relations: . David Magnusson

Chapters

Akita, Chiba, Fukui, Fukuoka, Gifu, Gunma, Hamamatsu, Himeji, Hiroshima, Hokkaido, Ibaraki, Iwate, Kagawa, Kagoshima, Kanazawa, Kitakyushu, Kobe, Kumamoto, Kyoto, Matsuyama, Miyazaki, Nagasaki, Nagoya, Nara, Niigata, Okayama, Okinawa, Omiya, Osaka, Sendai, Shinshu, Shizuoka, Tochigi, Tokushima, Tokyo, Toyohashi, West Tokyo, Yamagata, Yamaguchi, Yokohama.

Special Interest Groups

Bilingualism; College and University Educators; Computer Assisted Language Learning; Gender Awareness in Language Education; Global Issues in Language Education; Japanese as a Second Language; Junior and Senior High School Teaching; Learner Development; Material Writers; Other Language Educators; Pragmatics; Professionalism, Administration and Leadership in Education; Teacher Education; Teaching Children; Testing and Evaluation; Eikaiwa (Forming SIG); Pronunciation (Forming SIG); Teaching Elderly Learners (Forming SIG).

JALT Central Office

Urban Edge Building, 5F 1-37-9 Taito, Taito-ku, Tokyo 110-0016, Japan
Tel.: 03-3837-1630; Fax: 03-3837-1631; E-mail: jalt@gol.com;
Website: <www.jalt.org>

In this Issue

Articles

The main section of this issue contains three articles. **Jerry Greenfield** examines the issue of readability formulas for use in EFL. **Hideki Sakai** takes a look at the roles of output and feedback to enhance noticing by learners. **Emmanuel Manalo**, **Satomi Mizutani**, and **Julie Trafford** investigate the use of mnemonics to learn kanji.

Perspectives

Is Japanese English a distinct variety? **Philip Morrow** considers this and other questions from the perspective of World Englishes.

Reviews

This time we have a bumper crop of book reviews. The first four reviews look at methodology and teaching issues. **Thomas C. Anderson** reviews a teacher resource book on task-based learning, and **Roger Kenworthy** reviews an anthology of teaching methods. **Paul Hullah** examines a book on practical language teaching while **Michael J. Crawford** reviews a book on teacher professional development through narrative inquiry. The next two reviews are of books in the field of applied linguistics. **Scott Petersen** reviews a book on controversies in the field and **Scott Bronner** tackles a book on applied linguistics in language teaching. Computer-Assisted Language Learning (CALL) is the topic of the book reviewed by **Justin Charlebois**, and **Donna Tatsuki** reviews an introductory book on pragmatics. The review by **Jack Massalski** is of a book on bilingualism and the final review, by **Jonathan M.W. Rankine**, is of a book on contemporary Japanese culture.

From the Editors

With more books being published annually than ever before, the selecting of quality books to draw to the attention of readers has become an even greater challenge for the editors. We encourage you to write and contribute reviews of worthy books that you have encountered, especially reviews that look at two or three books on a related topic. We can thereby help one another stay well informed about the topics important to our profession.

The editors would like to welcome Deryn Verity to the *JALT Journal* Editorial Advisory Board. We appreciate the generous contribution of her time to help ensure the continued high quality of our publication.

Conference News

FEELTA Conference. The Far Eastern English Language Teachers Association (FEELTA) and The Institute of Foreign Languages of the Far Eastern State University will hold its third international conference on the problems of language and culture. The motto of the conference is: "People, Languages and Cultures in the Third Millennium." June 15-18, 2000 in Vladivostok, Russia. The aims of the conference are: To help transmit ideas and strategies of effective teachers and researchers; To consider the most vital problems of teaching culture and variants of the English language; To help establish professional connections between teachers and researchers of the Russian Far East and their overseas colleagues. For information contact Krainik Larissa, conference coordinator; postal address: Russia 690600, Vladivostok, Aleutskaya St., 56, Room 409; fax:(4232)25-72-00; e-mail: feltacon@dvgu.ru; or visit the homepage at <<http://www.dvgu.ru/rus/partner/education/feelta/>>.

The Fourth Symposium On Second Language Writing, Sept. 30 to Oct. 2, 2004, Purdue University, West Lafayette, Indiana, USA. Theme: "Second Language Writing Instruction in Context(s): The Effects of Institutional Policies and Politics." 20-minute presentations that address how instructional policies and politics affect instructional practices should include (1) a description of a particular L2 writing instruction context, (2) an analysis of how institutional policies and politics shape the curriculum in this context, and (3) a discussion of implications for second language writing theory, research, instruction, assessment and/or administration as well as the professional development of second language writing specialists. **Special Event: Graduate Student Conference on Second Language Writing.** This special event held in conjunction with the Symposium provides opportunities for graduate students to present their research and scholarship on second language writing to and receive feedback from their peers. Any topic related to second language writing or writing instruction is welcome. All presenters must be full-time graduate students. Proposals must be received by May 15, 2004. For more information, please visit: <http://symposium.jslw.org/2004/>. Tony Silva and Paul Kei Matsuda, Chairs.

Articles

Readability Formulas For EFL

Jerry Greenfield

Miyazaki International College

EFL/ESL teachers use English readability formulas to match texts to their students' reading levels. However, the formulas' validity for EFL/ESL use has gone largely untested. Two studies have now addressed this issue, with divergent results. Brown (1998) found that classic formulas were not very accurate predictors of EFL difficulty, while Greenfield (1999) found that they predicted for EFL about as well as they did for native English readers. Both studies produced accurate EFL readability formulas. In the analysis presented here, the difference in the two studies' findings is attributed to Brown's random passage set. Brown's formula proves more accurate with the other study's passages than with his own, agreeing with observed EFL difficulty and predictions by classic formulas. This supports the finding that the classic formulas are valid for EFL use.

EFL/ESL教師は英語読解難易度推定式を使って学生の読解力水準に合った英文を選定する。しかし、これまでこの推定式のEFL/ESLにおける有効性については殆ど検証されてこなかった。この問題については、現在までに二つの考察がなされたが、それぞれの結果は異なっている。ブラウン(1998)の考察では、標準的推定式によるEFL難易度の推定精度はあまり高くないとしているが、グリーンフィールド(1999)は、EFL及び母語読者双方に対するその読解力推定精度はほぼ同等であるとしている。これら二つの考察は、それぞれ正確なEFL読解難易度推定式を導き出している。二つの知見の違いは、ブラウンの無作為に選択した文節セットに帰因すると考えられる。ブラウンの推定式の精度は、自らが選んだ文節セットに対してよりも、グリーンフィールドの考察で使われた文節に対しての方が精度が高く、またそれは実際のEFL難易度と標準的推定式から得られる推定結果とも一致している。そしてこのことは、標準的推定式のEFLに対する有効性を裏付けるものである。

EFL/ESL teachers along with other English teachers have long turned to readability formulas for aid in matching texts to students' reading levels. Until recently little attention has been paid to whether it is appropriate to apply these tools outside the native English contexts in which the formulas were originally developed. This question has now begun to be investigated, but with mixed results. A study by J. D. Brown (1998) found that several classic formulas were not very accurate in predicting EFL reading difficulty for Japanese university students. His conclusions cast serious doubt on the validity of the classic readability formulas for EFL use. Instead Brown proposed a new formula of his own that he found to be more accurate with his readers. Unfortunately, that formula is difficult to use, requiring long-word and passage-frequency word counts in addition to parsing into function and non-function words. Brown's results thus leave EFL teachers without an easy and reliable way to estimate a text's readability for their students.

At the time Brown's article appeared, another study (Greenfield, 1999) with another group of Japanese EFL readers was being completed. It found that the classic formulas discriminated text difficulty for those readers about as well as they do for native English readers. The study produced a formula scaled to those Japanese readers which is easy to use but offers only a marginal improvement in predictive accuracy over the traditional formulas, which themselves proved to be quite satisfactory in the EFL as well as native speaker contexts. While the results of this study are encouraging, they appear to disagree with Brown's on the question of the fundamental validity of applying the classic formulas in EFL contexts. Where does this leave us regarding this question and the viability of the proposed EFL formulas? It will be useful to compare the two studies more closely.

Readability Formulas

Very simply, readability formulas are multiple regression equations in which the dependent variable (the value we want to know) is the reading difficulty predicted of a text and the independent or predictor variables are two or more directly measurable characteristics of the text, such as the number of letters per word and the number of words per sentence. To use one of these formulas, you measure the independent variables in a piece of text, plug those values into the formula, do the math, and get a prediction of the text's difficulty expressed as a grade level, a cloze score, or a score on some set scale. Dozens of formulas

have been introduced, and some of the most popular formulas, such as the Flesch (1948) and Dale-Chall (Chall & Dale, 1948), have been around since the 1940s (for an overview of readability formulas and their histories see Chall, 1958, 1988; Klare, 1963, 1974-75, 1988). For a time during the 1980s, readability formulas came under attack because of their low face validity when viewed from the vantage point of psycholinguistic theories of reading (Bruce & Rubin, 1988; Rubin, 1985; Bruce, Rubin, & Starr, 1981; Smith, 1988). Nevertheless, the formulas have survived and are still widely used on account of their consistently high predictive validity (Chall & Dale, 1995; Fry, 1989). That is to say, they have been found empirically to do a good job of discriminating text difficulty even though it is not obvious why they should or how they could. In fact, the advent of computer word processing has made the formulas more accessible than ever. In older versions of Microsoft Word, it was possible to get a readability report on a Word document using the Flesch, Flesch-Kincaid, Coleman-Liau, and Bormuth formulas built into that application. With Microsoft Word 97/98, the number of formulas was cut back to include only the Flesch and Flesch-Kincaid, but they are there waiting to be used at the click of a mouse. The question is, are they valid to use for EFL/ESL?

Although there are many different variables that have been identified as playing a part in reading difficulty (Gray & Leary, 1935), factor analysis has narrowed these down to only a few which have high correlations with the others and so can be used to represent them. The predictor variables in classic readability formulas typically represent just two main text factors: vocabulary difficulty and grammatical difficulty. Depending on the formula, vocabulary difficulty may be represented as word familiarity, average word length in syllables, proportion of long words, average word length in either characters or syllables, or proportion of monosyllable words. Grammatical difficulty is typically measured by the average number of words or syllables per sentence, based on a strong association of sentence length with, for example, the incidence of compound-sentence and embedded-clause constructions, which are much harder to count. Proposed approaches to readability measurement using variables that are not so easily countable have not been widely adopted. Nor do they need to be, since research has found them not to deliver significantly better results than formulas with simpler variables (Bormuth, 1969, 1971; Chall & Dale, 1995).

Generally the predictive accuracy of the most commonly used formulas has been found to be very high, yielding correlations with inde-

pendent comprehension tests in the .8 or .9 range (Chall, 1958; Chall & Dale, 1995; Fry, 1989). Readability formulas assume that the readers for whom they predict difficulty and the texts to which they are applied are similar to the samples used to derive the formulas in the first place. Indeed, almost without exception the formulas have been validated by testing American native English readers. Using the formulas to predict difficulty for second language readers assumes that those readers are not significantly different from native readers in ways that affect how the measured text variables relate to reading difficulty. Surprisingly, this assumption has been left essentially untested.

The question might be pursued on theoretical grounds, if we had a detailed enough model of second language reading. However, the theoretical issues in this area are many and complex, and empirical studies to investigate them have not yet produced a clear and comprehensive account of second language reading and how it is similar to or different from first language reading. (For an overview of the state of the research up to 1990, see Bernhardt, 1991; see also Grabe, 1993; Paran, 1996. For a sampling of recent studies, see Carrell & Wise, 1998; Parry, 1996.) A new survey is overdue. In any case, arguing from a theoretical model can only make the case for or against whether native reader-based formulas *ought* to work with second language readers. Such arguments cannot establish in fact whether they *do* work.

Validating Formulas for EFL

Fortunately, it is not necessary to resolve the theoretical issues in order to determine whether the formulas historically based on L1 reading data are also valid for EFL/ESL readers. Readability formulas represent statistical correlations and *predict* difficulty rather than *explain* its causes. Formula validity depends simply on the accuracy of predictions. This can be determined for EFL/ESL readers empirically by testing them to see how closely their performance matches what formulas predict. The remarkable fact is that this was not done to settle the matter years ago.

Hamsik's Study

In fact, a small-scale study was done in 1984 by Hamsik, who investigated the ESL validity of the Flesch, Dale-Chall, Fry, and Lorge readability measures. Hamsik gave cloze tests on 18 academic passages to 40 Intensive English Center students at an American university. The students

are described as being from the Middle East, South America, and “the Far East.” Hamsik found significant positive correlations of .775 to .819 between the rank orders of difficulty of the passages as indicated by the cloze scores and as predicted by each of the four readability measures. On the strength of this evidence, Hamsik concluded that “the four readability formulas and graphs...do measure readability [for] ESL students and that they can be used to select material appropriate to the reading level of ESL students” (p. iv).

Hamsik’s small heterogeneous sample of ESL readers did not permit discriminating any effect of first language background. With this in mind, Hamsik included among her recommendations one that “future studies of this sort should take account of L1 background” (p. 55). She also suggested that it might be possible to develop a readability index for ESL students that would be more accurate than existing formulas.

Brown’s Study

Further investigation of ESL/EFL formula validity was not forthcoming until Brown’s 1998 article. In an earlier study of cloze item difficulty, Brown (1992) had administered cloze tests to nearly 2300 Japanese EFL university students. For the new study he reanalyzed the data for difficulty at the passage level and compared the observed mean cloze scores on the passages with scores predicted by six readability measures: the Flesch, Flesch-Kincaid, Fry Graph, Gunning, Fog Count, and Gunning-Fog. Brown found Pearson correlations ranging only from .48 to .55, leading him to conclude, “first language readability indices are not very highly related to the EFL difficulty” (p. 27).

To address this need, Brown developed a new formula using his observed EFL scores as the criterion, scaled to yield an EFL Difficulty Index ranging from 1 to 92. Multiple regression analysis found the best fit or most accurate prediction to be made using four text variables: syllables per sentence, passage frequency (how many times the deleted item appears elsewhere in the text), percentage of long words (seven or more letters), and the percentage of function words. The resulting formula, which he called the EFL Difficulty Estimate, had a multiple correlation of .74, which yielded an adjusted R-Square or coefficient of determination of .51.¹ Usually a coefficient in that range would not be considered particularly strong. Nonetheless, because his formula yielded a stronger correlation with the observed EFL scores than did the classic formulas in his tests, Brown speculated, “EFL/ESL readability might best be esti-

mated separately for students from different language backgrounds" (p. 30). In other words, Brown suggests, we need to replace the classic readability formulas with new formulas specific to different language groups. His formula was offered as one that might be used with Japanese EFL. Brown's formula is as follows:

Brown EFL Difficulty Estimate

$$\begin{aligned} \text{EFL Difficulty} = & 38.7469 + (.7823 \times \text{Syllables per Sentence}) \\ & + (-126.1770 \times \text{Passage Frequency}) \\ & + (1.2878 \times \% \text{ Long Words}) \\ & + (.7596 \times \% \text{ Function Words}) \\ (R = .74, \text{adjusted } R^2 = .51, SE = 19.68, N = 50, p < .00001) \end{aligned}$$

Again, the formula is scaled to predict passage difficulty from zero to a maximum difficulty of 100. Something is amiss, however, because when this formula was applied to Bormuth's standard passage set (Bormuth, 1971) assembled for that researcher's testing and used again to calibrate the New Dale-Chall readability formula, Brown's formula predicted difficulty scores ranging from -276 to +64.

The Miyazaki Study

The Miyazaki study (Greenfield, 1999) also involved Japanese university students and checked the Flesch Reading Ease and Flesch-Kincaid formulas along with the Coleman-Liau, New Dale-Chall, and Bormuth formulas. The EFL participants in this study were 200 Japanese students enrolled in a small liberal arts college in western Japan. Careful randomized testing procedures were followed, based on Bormuth (1971). Fifth-word deletion cloze tests were constructed on 31 of the 32 Bormuth academic passages. One passage was read by all participants as a control, and one was omitted for a balanced design. Pearson correlations between observed EFL mean cloze scores and scores predicted by the formulas are .691 for the New Dale-Chall formula, .765 for Coleman-Liau, .845 for Flesch Reading Ease, .847 for Flesch-Kincaid, and .861 for Bormuth. These results are shown in Table 1.

Bormuth left us his set of mean cloze scores for his 1971 passages, so it is possible to compare the observed EFL scores and Bormuth's native English reader criterion for the same passages.² That correlation is even stronger at .915. These correlations are generally consistent with

Table 1. Pearson Correlations Between Scores for Bormuth Passages Predicted by Miyazaki EFL Index, Original and Recalculated Brown Formulas, and Classic Formulas

	Observed EFL	Miyazaki EFL Index	Brown (Original)	Brown Recalculated
Flesch	-.845	.957	-.764	.902
Flesch-Kincaid	-.847	.980	.775	-.909
Coleman-Liau	-.765	.927	.736	-.850
Bormuth 1969	.861	.977	-.775	.913
Dale-Chall 1995	.691	.790	.654	-.820
Brown Recalculated	.907	.825	-.926	
Bormuth 1971	.915	.944	-.781	.890
Observed		.941	-.820	.895

All relationships significant at $p < .0001$, $N = 31$. The negative sign can be ignored as an artifact of contrasting scales.

inter-formula correlations and predictive accuracies reported in the L1 literature (e.g. Chall & Dale, 1995). These findings support the conclusion that the classic formulas are indeed fundamentally valid for a broad spectrum of English readers that includes non-native as well as native readers. In other words, the formulas work quite well to predict the relative EFL/ESL difficulty of English academic texts.

What remained was to check whether recalculating the classic formulas using the EFL scores would significantly improve accuracy with EFL readers. While some of the recalculations result in a small but statistically significant improvement, the gain did not seem sufficient to justify substituting them for the originals. Going a step further, a comprehensive check of all of the classic variables found that a regression of just two, letters per word and words per sentence, against the study's EFL criterion produced an EFL difficulty index that was as good as or slightly better than any of the classic formulas. The ANOVA and coefficients for that multiple regression are shown in Table 2.

The new formula, for convenience called the Miyazaki EFL Readability Index, turned out to be only marginally more accurate than the classic formulas. However, like Brown's, it has the practical advantage of

Table 2. ANOVA and Regression of Brown's Four Independent Variables Forced vs. Bormuth EFL Criterion

	DF	Sum of Squares	Mean Square	F Value	p-Value
Regression	4	4093.682	1023.420	29.990	< .0001
Residual	26	887.272	34.126		
Total	30	4980.954			
	Coefficient	Std. Error	Std. Coeff.	t-Value	p-Value
Intercept	33.232	9.649	33.232	3.444	.0020
Syllables per Sentence	-.249	.205	-.177	-1.216	.2351
Passage Frequency	12.834	3.045	.486	4.215	.0003
Long Words	-48.665	21.996	-.327	-2.212	.0359
Function Words	-65.650	26.595	.225	-2.468	.0205

N = 31

being scaled for EFL readers while, unlike Brown's, being simple to apply using easily found word counts. The Miyazaki formula is as follows:

Miyazaki EFL Readability Index

$$\begin{aligned} \text{EFL Difficulty} = & 164.935 - (18.792 \times \text{Letters per Word}) \\ & - (1.916 \times \text{Words per Sentence}) \\ (R = .862, \text{ adjusted } R^2 = .723, SE = 10.558, N = 31, p < .0001) \end{aligned}$$

Note that this formula delivers a reading ease score on a nominal 100-point scale, 100 being easiest.

Comparison of Brown's and the Miyazaki Study

On the face of it, Brown's study and the Miyazaki study seem to have come to contrary conclusions regarding the validity of classic readabil-

ity formulas for EFL use, leaving the matter still unresolved. However, a closer look reveals that the disagreement is less direct than it seems at first glance. In spite of obvious parallels in the research designs, there are important differences in what the two studies were actually testing.

Participants

Let us look first at the population samples. It seems reasonable to assume that the groups participating in the two studies were similar to each other except with respect to their size. Brown's group was comprised of 2,298 participants distributed across 18 universities in Japan, while the 200 participants in the Miyazaki study were from a single college. While ordinarily the larger sample is advantageous for statistical purposes, in this case there is no reason to suppose that the Miyazaki group was in fact importantly different from the group in Brown's study. The two population samples were each internally homogeneous and were similar to each other in first language, cultural background, and general educational level.

Brown tells us little about the English proficiency of his group beyond the fact that they were representative of Japanese university EFL students in this regard. A similar claim is made for the Miyazaki students. However, even if the two groups had differed in their English proficiency levels, there is no reason to suppose that the Miyazaki group was higher than Brown's rather than the other way around. In any case, it is the difference in accuracy rates obtained in the two studies—Brown's group scoring much lower overall than the Miyazaki group—that begs to be explained. More importantly it must be wondered why there is such a large difference in correlations with predictions by the classic formulas. On balance it seems very unlikely that a difference in the English proficiency of the two EFL population samples or any other differences in the two population samples could explain the large differences in the correlations between formula predictions and observed difficulty found in the two studies.

Passage Sets

The more likely source of the disagreement between the two studies lies in differences in the nature of their passage sets arising in turn from difference in the purposes of the studies. The Miyazaki study was looking at whether the classic formulas' ability to discriminate relative

difficulty is different for EFL and native English readers. To answer this question, it was desirable to compare EFL performance with native English reader performance while keeping the passage set constant. Using Bormuth's passages made it possible to compare observed EFL difficulty with his observed native English reader scores as well as with the New Dale-Chall formula predictions based on the same passages. Using the Bormuth passages with an EFL group focused squarely on whether the text variables *for those texts* relate to EFL difficulty the way they do for native English readers. The issue of the formulas' applicability to other kinds of texts was not addressed.

Brown, on the other hand, chose not to control for difference in passage type but was checking the universality of the classic formulas for predicting the difficulty of randomly selected texts. His method of selecting passages was quite different from that followed by Bormuth and other classic readability researchers. Typically a criterion passage set is deliberately chosen to exhibit a well-distributed range in the values of text variables to be regressed against test scores. Instead, Brown's passages were randomly selected to be "representative samples of the English language, at least the English language written in the books found in a U.S. public library" (Brown, 1998, p. 16). Brown verified their representativeness by comparing their lexical frequencies with frequencies published for the English corpus, finding them to correlate at .93. The ostensible advantage of such an approach is that it tests the EFL validity of the classic formulas not for only *academic* materials but rather for *any* English texts. To suppose that such a set is superior is to suppose that the academic materials traditionally used may not be fully representative of English texts at large. In effect, Brown concluded that *when applied to more generally representative texts than the kinds of academic texts on which they are based*, classic readability formulas do not work very well for EFL readers. Thus, Brown's finding does not directly contradict that of the Miyazaki study on the more specific question as to whether the formulas might, however, be valid for EFL readers when applied to academic texts.

Brown's conclusion is complicated by the fact that he was concerned with the issue of text representativeness as well with the question of whether the first language of the readers makes a difference in formula validity. As a result, it is not clear whether the inability of the formulas to predict the scores he observed was due to the fact that his readers were second language readers or to the fact that his texts were randomly selected, or some combination of both. Brown's suggestion that different formulas need to be developed for different EFL/ESL language

backgrounds implies that he believes that language background was the more important difference. However, since no one (to my knowledge) has ever provided a native reader difficulty criterion based on a random passage set, there is nothing against which to compare the EFL results to answer that question, and so Brown's data must be regarded as ultimately inconclusive on this point.

It might be argued that using a random passage set corrects for a shortcoming in the classic tradition of readability research, namely that that tradition has focused on a too narrow range of text types and, thus, has failed to achieve full generalizability. Such an argument is unconvincing on two counts. First, the principal use of readability formulas generally is in education-related contexts, where they are used not only by teachers but even more importantly by writers and publishers of educational materials. No claim is made that the various criterion passage sets used to derive the classic formulas are or should be representative of the English language in general. In fact, Bormuth deliberately selected his passages to represent the range of content, structure, style, and vocabulary specifically found in school texts. The decision to select criterion passages to represent educational materials in this way does not, on that account, limit the practical usefulness of formulas derived from them. On the contrary, there is little point in creating readability formulas for other kinds of texts. This is especially so for formulas to be used in EFL contexts.

Second, as Brown himself had already pointed out in his earlier article on natural cloze tests (Brown 1993), there is no assurance that a randomly selected set of texts will provide the variability in the text features needed to discriminate difficulty, regardless of how representative their combined lexical frequencies may be. In fact, Brown's data suggest that the passages did not work especially well for this purpose. The range of mean cloze scores for the passages was relatively flat and the passage scores overall very low (mean = 13.7%), with fully 40% of the mean passage scores falling below 10% accuracy. Brown attributes these low accuracy rates to the nature of cloze tests. However, in the Miyazaki study only 10% of the mean passage scores fell below 10% accuracy, with the range of mean raw scores much greater and the mean score for the set also considerably higher at 24.25%. The evidence suggests that Brown's randomly selected passages may have been too difficult overall to yield a robust difficulty variable to compare with formula predictions. To make the scores more suitable for regression analysis, Brown scaled them in such a way that the variability was statistically magnified. Although this

is a perfectly reasonable procedure, it may explain the out-of-range results found when his resulting formula is applied to a passage set more variable than his random set apparently was.

Brown's EFL Difficulty Estimate

This brings us back to Brown's formula. Although it correlates only weakly with the classic formulas, Brown's EFL Difficulty Estimate nonetheless is moderately strong in discriminating the readability of his passages for his representative sample of Japanese EFL students. But there apparently were problems with the passage set, and anyway it is unknown how that set relates to typical academic texts. We might therefore ask how the formula works for academic texts as represented, for example, by Bormuth's passages. This question was answered by applying the EFL Difficulty Estimate to those passages and comparing the results with predictions by the classic formulas, with Bormuth's native reader scores, and with the Miyazaki EFL scores on the passages. Those correlations are included in Table 1.

The results are a little surprising. The correlations between scores predicted for the Bormuth passages by classic formulas and scores predicted by Brown's formula range from .654 to .832, with Flesch at .766, Flesch-Kincaid at .778, and only the correlation with New Dale-Chall falling below .734. These are much higher than Brown found with his own passages. When Brown's estimates for the Bormuth passages were compared with the observed Miyazaki EFL mean cloze scores, the resulting correlation of .841 was comfortably consistent with correlations ranging from .691 to .861 found between the observed EFL difficulty and classic formula scores for those passages. In other words, Brown's formula worked about as well on the Bormuth passages read by the Miyazaki students as do the classic formulas and the Miyazaki formula.

It remained only to see whether Brown's formula could be improved further by recalculating its coefficients using the Miyazaki scores. This involved performing a new multiple regression using the four variables used in Brown's model. The ANOVA and coefficients for the four-variable regression are shown in Table 3.

Normally when performing a multiple regression, variables are entered or removed one at a time to discover which combination results in the strongest multiple correlation, or best fit, with the fewest variables. Since in this case all four variables are being forced into the equation, there is a possibility that one or more variables might be extra bag-

Table 3. ANOVA and Regression of Two Independent Variables vs. Miyazaki EFL Criterion (Miyazaki EFL Readability Index)

	DF	Sum of Squares	Mean Square	F Value	p-Value
Regression	2	8988.565	4494.282	32.874	< .0001
Residual	28	3121.061	111.466		
Total	30	12109.626			

	Coefficient	Std. Error	t-Value	p-Value
Intercept	164.935	18.758	8.793	<.0001
Letters per Word	-18.792	4.940	-3.803	.0007
Words per Sentence	-1.916	.484	-3.793	<.0005

*N = 31

gage, adding nothing to the strength of the regression. In fact, the three word-frequency variables were well within bounds, but the syllables-per-sentence variable failed the relatively liberal .1 probability limit to add/remove adopted by Brown in his (more proper) stepwise regression procedure. Removing this variable improved the probability of the remaining three variables and only very slightly reduced the adjusted R Squared from .794 to .791, which is very strong. The four-variable recalculated formula is as follows:

Recalculated Brown EFL Difficulty Estimate

$$\begin{aligned} \text{Cloze} = & 33.232 + (-.249 \times \text{Syllables per Sentence}) \\ & + (12.834 \times \text{Passage Frequency}) \\ & + (-48.665 \times \% \text{ Long Words}) \\ & + (-65.650 \times \% \text{ Function Words}) \end{aligned}$$

($R = .907$, adjusted $R^2 = .794$, $SE = 5.842$, $N = 50$, $p < .0001$)

Note that EFL difficulty in the recalculated formula is given as a predicted cloze score and is not rescaled as in Brown's original formula.

While the three-variable version is slightly more accurate with the sentence-length variable dropped out, it has no variable ostensibly rep-

representing a syntactic factor. The function-word variable, which might seem to be related to syntax, in fact is correlated in these passages only .174 with syllables per sentence, which is ordinarily taken to be a syntax variable and in this case is itself correlated .766 with observed difficulty. Since predictions by the two versions are nearly perfectly related at .994, it is redundant but otherwise does no harm to retain the four-variable recalculation as the more faithful to Brown's original model. Note that this recalculated formula has an adjusted R Squared of .794, much higher than the .51 achieved by Brown's original regression.

We may then compare predictions of the recalculated Brown formula with the observed EFL scores, Bormuth's 1971 scores, predictions of the original Brown formula, and predictions of each of the classic formulas. These correlations are also included in Table 1. In some of the corresponding correlations it is readily apparent that one correlation is stronger than its counterpart. At the same time, it is not obvious, particularly in pairs, which are closer in value, whether a given difference is significant. There is a test, the Williams *t*-test, that finds whether a difference between two related correlations is significant.³ In this case it allows us to determine whether the Miyazaki formula, the original Brown formula, and the recalculated Brown formula are equally good, better, or worse as predictors of cloze performance in comparison with each other and with each of the classic formulas.

The results of the Williams *t*-test are shown in Table 4. The recalculated Brown formula has a small but statistically significant advantage over Brown's original formula, and both are superior to all of the classic indices except the Bormuth formula. The comparison shows the Miyazaki Index to correlate more strongly than Brown's formula, but the difference was not statistically significant.

The bottom line is that we have two new formulas developed from EFL data that appear to work very well in predicting the relative difficulty of academic texts. At the same time we have strong evidence that the new formulas have only a narrow, if any, advantage over the time-tested traditional formulas, especially the Flesch and Flesch-Kincaid, and Bormuth formulas. We may therefore use those formulas with some new confidence that they are valid for EFL. By extension, if they are valid for a first language group as different from English as Japanese, they are probably valid for other EFL and ESL contexts as well. Brown's own findings do not directly contradict this conclusion, if we understand that his passage set is not strictly comparable and may not have been appropriate for basing a measure of academic readability.

Table 4. Williams *t*-test: Recalculated Brown Formula versus Classic, Original Brown & Miyazaki EFL Formulas

Named Formula	Recalculated Brown vs Observed EFL	Named Formula vs Observed EFL	Named Formula vs Recalculated Brown	Williams <i>t</i> -value
Flesch	<u>.907</u>	-.845	.862	4.061
Flesch-Kincaid	<u>.907</u>	-.847	.864	4.005
Coleman-Liau	<u>.907</u>	-.765	.820	6.777
Bormuth 1969	.907	.861	.807	1.893
Dale-Chall 1995	<u>.907</u>	.691	.591	3.939
Original Brown	<u>.907</u>	.820	.915	8.573
Miyazaki EFL Readability Index	.907	.941	.808	2.462

Underlined values are significantly larger in that row's comparison at $p < .01$, with $t = 2.762$ needed for significance (2-tailed, $df = 28$). The negative sign can be ignored as an artifact of contrasting scales.

Applications

If the old formulas are already valid, is there any point in introducing a new formula that is not significantly or importantly more accurate? The answer is not as simple as it might seem. Predictive accuracy is only part of the story. Along with accuracy, we need to consider how easy any index is to apply. Bormuth took account of this when he constructed different formulas for hand scoring, computer calculation, and unrestricted research use. The Coleman-Liau formula (Coleman & Liau, 1975) is specifically intended for computer calculation. For application in the field, the ease of a simpler hand scoring formula compensates for a small loss in accuracy. Part of the attractiveness of the Flesch and Flesch-Kincaid (Kincaid, Fishburne, Rogers, & Chissom, 1975) formulas has been their ability to deliver accurate results with just two simple variables that are easy to count and calculate. Of course, this advantage disappears with computerized applications. However, it is not realistic to expect we would ever have an EFL formula built into Microsoft Word. Fortunately, the Miyazaki results indicate that the Flesch and Flesch-Kincaid formulas included in that application already can serve our needs.

Viewed against this availability, the complexity and difficult-to-count variables of Brown's formula make hand scoring with that formula not very attractive. The Miyazaki formula uses only two variables that not only are easy to count but are reported along with the readability report in Microsoft Word. Other word-processing applications with or without readability measurement also provide character, word, and sentence counts, making hand scoring any passage a simple affair. To make it even easier to use the Miyazaki index, the Miyazaki study provides a lookup table of scores (Greenfield, 2003) for a practical range of word- and sentence-length values, so no calculation is needed. Once you have these values, whether by counting or by using a word processor, it is straightforward to locate a score on the table that represents the EFL readability of the passage. The score is figured on a 100-point scale, with 100 being easiest and 50 representing a text of average difficulty for EFL students.

What these scores actually mean, however, is not so straightforward. EFL students achieved a mean cloze accuracy of about 27% on a text having a Miyazaki EFL Readability Index of 50. No one knows for sure what that means in terms of performance criteria for EFL reading. The conventional wisdom for native readers is that a score of 35% corresponds to a score of 50 on a well-constructed multiple-choice comprehension test, 45% corresponds to an MC score of 75, and 55% to an MC score of 90. Alternatively, some have said a 35% cloze accuracy rate is satisfactory for a text to be read with classroom support, 45% for homework, and 55% for extensive reading for pleasure. These suggested figures are open to question, however, even for native readers.

It might be thought that EFL readers have a lower tolerance for texts that are challenging than native readers do, but I suspect that the opposite is true. Applying a cloze accuracy criterion of 45% to texts used regularly in my own institution's content-based EFL classes would find many of them out of reach for most of the students who nonetheless do successfully read them. If this is so, is it because EFL students are prepared to work harder to comprehend a text than are native English students? Do EFL students have a higher tolerance for imperfect decoding and use other strategies to comprehend the text? Do EFL readers comprehend texts in a general way better than they are able to produce accurate cloze completions of individual items? In general, does the relationship of cloze accuracy to general comprehension tend to be different for EFL readers than for native readers?

That research has yet to be done. The problem of establishing performance criteria for EFL readability is still very much unresolved (Greenfield, 2001, 2003). Finding that the formulas do a good job of discriminating relative difficulty does not by itself address this question. Neither the Miyazaki study nor Brown's has taken this on as a research issue. Interpreting formula results to help make a determination about what reading materials are appropriate for a particular context still calls on the expertise of the successful teacher. We can now be more confident, however, that the information delivered by the formulas is in fact reliable and relevant to making such a judgment.

Jerry Greenfield is Professor of English at Miyazaki International College, where he also teaches courses in Information Technology and Aesthetics. In addition to his research interests in second language readability and reading assessment, he has presented and published on issues in Web page design and legibility.

Notes

1. This is the present author's calculation for comparison with the Miyazaki results; Brown reports the unadjusted R Square of .55. The adjusted coefficient takes account of the number of variables and provides a more precise estimate of validity.
2. Bormuth gives these scores in a probit metric to remove certain floor and ceiling effects in his data. This was accomplished by looking up percentage scores in a table showing the area under a normal distribution curve to find the corresponding deviation score. He then scaled the scores in such a way that cloze scores of .10, .30, and .50 took on probit values respectively of 372, 448, and 500 (Bormuth, 1971, p. 88). Without Bormuth's original raw scores, it is impossible to compare accuracies directly. However, this does not prevent using Bormuth's scaled criterion in testing correlations with other score sets for the passages.
3. Howell (1997). This test finds whether a difference between two related correlations is significant, or in this case whether a formula is equally good, better, or worse as a predictor of cloze performance in comparison with another formula. The formula for this test is as follows:

$$t = (r_{1,2} + r_{1,3}) \sqrt{\frac{(N-1) \cdot (1 + r_{2,3})}{2 \left(\frac{N-1}{N-3}\right) \cdot |R| \cdot \left(\frac{(r_{1,3} + r_{2,3})^2}{4}\right) \cdot (1 - r_{2,3})^3}}$$

$$\text{where } R = (1 - r_{1,2}^2 - r_{1,3}^2 - r_{2,3}^2) + (2r_{1,2} \cdot r_{1,3} \cdot r_{2,3})$$

In this equation $r_{1,2}$ is the correlation between observed cloze mean scores and one formula's scores, $r_{1,3}$ is the correlation between observed cloze mean scores and another formula's scores, $r_{2,3}$ is the correlation between the two formulas' scores, and N is the number of passages. This ratio is distributed as t on $N-3$ degrees of freedom.

References

- Bernhardt, E. B. (1991). *Reading development in a second language: Theoretical, empirical, and classroom perspectives*. Norwood, NJ: Ablex.
- Bormuth, J. R. (1969). *Development of readability analyses* (Final Report, Project No. 7-0052, Contract No. 1, OEC-3-7-070052-0326). Washington, DC: U.S. Office of Education.
- Bormuth, J. R. (1971). *Development of standards of readability: Toward a rational criterion of passage performance*. U. S. Department of Health, Education, & Welfare (ERIC Doc. No. ED O54 233).
- Brown, J. D. (1992). What text characteristics predict human performance on cloze test items? In *Proceedings of the 3rd conference on second language research in Japan, Tokyo, November 16, 1991* (pp. 1-26). Niigata: Language Programs of the International University of Japan.
- Brown, J. D. (1993). What are the characteristics of natural cloze tests? *Language Testing*, 10, 93-116.
- Brown, J. D. (1998). An EFL readability index. *JALT Journal*, 29, 7-36.
- Bruce, B., & Rubin, A. (1988). Readability formulas: Matching tool and task. In A. Davison & G. M. Green (Eds.), *Linguistic complexity and text comprehension: Readability issues reconsidered* (pp. 5-22). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bruce, B., Rubin, A., & Starr, K. (1981). *Why readability formulas fail*. Reading Education Report No. 28. Urbana, IL: University of Illinois Center for the Study of Reading. (ERIC Doc. No. ED 205 915)
- Carrell, P. L., & Wise, T. E. (1998). The relationship between prior knowledge and topic interest in second language reading. *Studies in Second Language Acquisition*, 20, 285-305.

- Chall, J. (1958). *Readability: An appraisal of research and application*. Columbus, OH: Ohio State University.
- Chall, J. (1988). The beginning years. In B. L. Zakaluk & S. J. Samuels (Eds.), *Readability: Its past, present, and future* (pp. 2-13). Newark, DE: International Reading Association.
- Chall, J., & Dale, E. (1948). A formula for predicting readability. *Educational Research Bulletin*, 27, 11-20.
- Chall, J., & Dale, E. (1995). *Readability revisited: The New Dale-Chall Readability Formula*. Cambridge, MA: Brookline Books.
- Coleman, M., & Liau, T. L. (1975). A computer readability formula designed for machine scoring. *Journal of Applied Psychology*, 60, 283-284.
- Dale, E., & Chall, J. (1948). A formula for predicting readability. *Educational Research Bulletin*, 27, 11-20, 28, & 37-54.
- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32, 221-233.
- Fry, E. (1989). Reading formulas: Maligned but valid. *Journal of Reading*, 32, 292-297.
- Grabe, W. (1993). Current developments in second language reading research. In S. Silberstein (Ed.), *State of the art TESOL essays: Celebrating 25 years of the discipline* (pp. 205-236). Alexandria, VA: Teachers of English to Speakers of Other Languages.
- Gray, W. S. and Leary, B. (1935). *What makes a book readable?* Chicago: University of Chicago Press.
- Greenfield, G. (1999). *Classic readability formulas in an EFL context: Are they valid for Japanese speakers?* Ed.D. dissertation, Temple University. University Microfilms No. 99-38670.
- Greenfield, G. (2001). Performance criteria for EFL/ESL reading difficulty. Paper presented at the TESOL 2001 Conference, St. Louis, Missouri, March 2, 2001 (Event #5271, audiotaped).
- Greenfield, G. (2003). The Miyazaki EFL Readability Index. *Comparative Culture*, 9, 41-49. Miyazaki, Japan: Miyazaki International College.
- Hamsik, M. J. (1984). *Reading, readability, and the ESL reader*. Unpublished doctoral dissertation, University of South Florida.
- Howell, D. C. (1997). *Statistical methods for psychology* (4th ed.). Belmont, CA: Duxbury Press.
- Kincaid, J. P., Fishburne, R. P., Jr., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel*. Research Branch Report 8-75. Millington, TN: Naval Technical Training, U. S. Naval Air Station, Memphis.
- Klare, G.R. (1963). *The measurement of readability*. Ames, IA: Iowa State University Press.

- Klare, G.R. (1974-75). Assessing readability. *Reading Research Quarterly*, 10, 62-102.
- Klare, G. R. (1988). The formative years. In B. L. Zakaluk & S. J. Samuels (Eds.), *Readability: Its past, present, and future* (pp. 14-34). Newark, DE: International Reading Association.
- Paran, A. (1996). Reading in EFL: Facts and fictions. *ELT Journal*, 50, 25-42.
- Parry, K. (1996). Culture, literacy, and L2 reading. *TESOL Quarterly*, 30, 665-692.
- Rubin, A. (1985). How useful are readability formulas? In J. Osborn, P. T. Wilson, & R. C. Anderson (Eds.), *Reading education: Foundations for a literate America* (pp. 61-77). Lexington, MA: Lexington Books (D. C. Heath).
- Smith, F. (1988). *Understanding reading: A psycholinguistic analysis of reading and learning to read*. Hillsdale, NJ: Lawrence Erlbaum Associates.

(Received June 5, 2003; revised November 30, 2003)

Roles of Output and Feedback for L2 Learners' Noticing

Hideki Sakai

Shinshu University

This study explores the roles of output and feedback in L2 learners' noticing and repair in later production. Sixteen Japanese learners of English, assigned to one of two conditions (model or recast), performed communicative tasks and afterwards reported what they noticed about errors or linguistic problems in a retrospective interview. Results showed that although the learners noticed linguistic problems by producing the target language, they had difficulty in attending to and incorporating the subsequent input for later production. Also the results indicated that recasts were more effective to prompt the noticing of errors than were models, and that noticing a *gap* through feedback led to a relatively higher number of repairs on the second trial than was the case of noticing a *hole* through output.

本研究は、アウトプットとフィードバックの役割を、気づきと学習効果（2度目の発話時における修正）の点から調べた。日本人英語学習者 16 名が、フィードバックとしてモデルが与えられたグループと言い直し (recasts) が与えられたグループに分けられた。学習者は、コミュニケーション・タスクの後で気づきについて言語報告した。主な結果は次の通りである。学習者は目標言語を表出することによって言えないことに気づいたが、その後でフィードバックによって与えられたインプットに注目して取り入れ、2度目の発話時に表出することは難しかった。また、言い直しはモデルよりも気づきを促進する点で効果的であった。さらに、フィードバックによってギャップに気づくこと (noticing a gap)の方がアウトプットによって言えないことに気づくこと (noticing a hole) よりも2度目の発話時の修正につながることを示された。

Noticing and Second Language Acquisition

The issue of noticing has drawn much attention in second language acquisition (SLA) research (e.g., Ellis, 1994; Gass, 1988, 1997; Long, 1991, 1996; Long & Robinson, 1998; Robinson, 1995; Schmidt, 1990, 1993, 1995; Schmidt & Frota, 1986; Skehan, 1998; Swain, 1985, 1998; Tomlin & Villa, 1994; Truscott, 1998). In his noticing hypothesis, Schmidt (1990, 1993, 1995; Schmidt & Frota, 1986) shed light on focal attention, or noticing, as a necessary and sufficient condition for input to become intake in SLA, by claiming "intake is that part of the input that the learner notices" (1990, p. 139). According to Schmidt (1993), second language (L2) learners need to not only comprehend the input but also notice "whatever features of the input are relevant for the target system" (p. 209).

An attempt has been made to categorize noticing based on what is noticed. Swain (1998, p. 66) pointed out that there can be different types of noticing such as: (a) noticing a form in the input, (b) noticing one's interlanguage deficiencies (or holes), and (c) noticing the gap between the interlanguage and the target language. The first type of noticing is when, while listening to or reading the target language input, the learner simply attends to the formal aspects of the target language in the input. Input frequency, saliency of the input, and external manipulations such as input enhancement, may influence this first type of noticing. In the second type, noticing the interlanguage deficiencies is when learners may notice that they cannot say what they want to say precisely in the target language. This type of noticing is also referred to as noticing "holes" (Doughty & Williams, 1998, p. 228; Swain, 1998, p. 66). Thus, the term *holes* has been adopted in this paper. Swain argued that one major function of production in the target language is to facilitate this type of noticing (1993, 1995). Finally, the third type of noticing is when learners may notice that their current interlanguage is different from the target language. Feedback provided during interactions may help learners notice their errors, that is, notice the gap (Long, 1996). The present study focuses on output and feedback because both are considered to promote different types of noticing, as will be seen in the following sections.

Output and Noticing

In her output hypothesis, Swain (1985, 1993, 1995) proposed that producing the target language contributes to SLA via several functions: (a) the automaticity function, (b) the hypothesis-testing function, (c) the

metalinguistic function, and (d) the noticing/triggering function (see also de Bot, 1996).

The last function of output is the most important to this study. Swain (1995) stated that “the activity of producing the target language may prompt second language learners to consciously recognize some of their linguistic problems; it may bring to their attention something they need to discover about their L2” (p. 126). Consequently, output enhances the likelihood of processing the subsequent input (Swain, 1993, p. 160). Swain and her colleague (Swain, 1998; Swain & Lapkin, 1998) showed that the activity of producing the target language helped L2 learners to notice their linguistic problems in the target language.

The question arises then: Does noticing the linguistic problems promote the noticing of relevant forms in subsequent input and result in learning? Izumi, Bigelow, Fujiwara, and Fearnow (1999) and Izumi and Bigelow (2000) attempted to answer this question, but found no empirical support for this assumption. They made two groups (an experimental group and a control group) and gave two types of writing tasks to both groups: text reconstruction tasks and guided essay-writing tasks. In the text reconstruction tasks, the experimental group read texts for the succeeding writing tasks; on the other hand, the control group read the same texts only for comprehension. In the guided essay-writing tasks, the experimental group performed writing tasks for given topics before the model input while the control group performed writing tasks for unrelated topics. Both groups wrote essays after the model input. The researchers examined the learners’ noticing while reading, the use of the target forms (the past hypothetical conditional) in their writing tasks, and the performance in the multiple-choice recognition tests and the picture-cued production tests. Results showed that the writing activity did not better enhance noticing of the target forms in the model input provided after the writing activity in the experimental group than in the control group. These studies did not lend support for Swain’s (1993, 1995) argument that noticing holes facilitates noticing the relevant forms in the subsequent input.

It should be noted that it was not clear whether their learners noticed their linguistic problems at the first production (for other methodological issues, see Whitlow, 2001; Izumi & Bigelow, 2001). As Izumi and Bigelow (2000) admitted, “not all learners necessarily found their IL [interlanguage] grammar to be problematic during production, which in turn affected their attention to the grammatical form when they were exposed to the input” (p. 271). Schmidt (1993) cautions that “the prob-

lem with this external approach is that the treatment may not have the intended effect” (pp. 218-219). Izumi and his colleagues assumed that writing tasks had caused their learners to notice their linguistic problems, but in fact they might have failed to notice their linguistic problems as to the target structure. Thus, it remains unclear how or even whether output functions as a facilitator of noticing.

Feedback and Noticing

In this section, I will focus on feedback and its effect on noticing. In his updated version of the interaction hypothesis, Long (1996) proposed that the negative feedback provided during negotiation for meaning may facilitate SLA because it effectively draws L2 learners’ attention to the linguistic forms in the course of interaction. Recasts are such feedback, defined as reformulation of L2 learners’ erroneous utterances by the interlocutor¹, maintaining the meanings which L2 learners intend to convey (Long, 1996, p. 434; Lyster, 1998a, p. 58).

Recasts occur most frequently in NS-NNS interactions (Oliver, 1995; Van den Branden, 1997) and in classrooms (Doughty, 1994; Lyster, 1998a, 1998b; Lyster & Ranta, 1997; Roberts, 1995). Also, several studies have attempted to show that recasts are effective for language learning² (Doughty & Varela, 1998; Long, Inagaki, & Ortega, 1998; Mackey & Philp, 1998; for a review, see Nicholas, Lightbown, & Spada, 2001). Recasts are considered to be effective for SLA in that they possess the dual function of providing positive evidence and negative evidence at the same time, which may prompt L2 learners’ noticing because it makes cognitive comparison easier (Long, 1996, p. 434). Here is an example from the present study.

(1) From the recast group

Hikaru: *His mother angry.*

Researcher: *His mother is angry?* [recast]

The learner, Hikaru³, produced an utterance that included an error related to the “Be” verb *is*. The recast was provided in the next turn in the form of a confirmation check.⁴ This feedback was a reformulation of the erroneous utterance and maintained the proposition intended by the learner. In this example, the recast indicated that (a) the omission of

the “Be” verb was wrong in English (negative evidence), and that (b) the correct form to be inserted was *is* (positive evidence).

As Ellis (1994, p. 96) pointed out, an external comparison appearing in the recast exchange may not always result in a cognitive comparison. The question then arises as to whether L2 learners actually notice the corrective nature of recasts. In other words, do recasts promote L2 learners’ noticing of the gap? The first evidence for this, though indirect, is that L2 learners are more likely to imitate recasts than they are to repeat a non-corrective repetition (Long, 1996, pp. 435-439). For example, Doughty (1994) found that university students learning French as a foreign language imitated 21.5% of the teacher’s recasts while they repeated 2.3% of the non-corrective repetition. Because recasts and non-corrective repetition have common functional distributions and are both repetitive in nature (Lyster, 1998a), L2 learners are considered to be responding to the corrective nature of recasts.⁵

Another approach to the question is to examine L2 learners’ noticing through introspection⁶ (Mackey, Gass, & McDonough, 2000). Mackey et al. (2000) utilized stimulated recall to explore how L2 learners perceive feedback in interactions. The participants for their study were 10 ESL learners and seven learners of Italian as a foreign language. The learners performed a communicative task, which was videotaped for a later recall session. They were then asked to report verbally what they had thought during the interaction. During the task, the feedback in the form of recasts and negotiation was naturally provided. Results of the analysis of the feedback and the verbal reports showed that morphosyntactic feedback was generally provided in the form of recasts and was not perceived as being about morphosyntax. In contrast, phonological and lexical feedback, generally in the form of negotiation and combination (negotiation and recast), was perceived more accurately by the learners. The researchers went on to state that “the window of opportunity for these learners to notice grammar in interaction may have been relatively small” (p. 488). Their study is of great interest because their findings suggest the possibility that L2 learners would not notice the target of the feedback as intended by the feedback provider. However, only a small number of studies have attempted to explore the noticing issue in order to investigate the roles of feedback. Further studies need to be carried out.

Purpose of the Study

The main purpose of this study was to examine the verbal reports and to clarify how L2 learners notice errors or linguistic problems in their production. The research questions posited for this study are as follows:

1. Does output promote L2 learners' noticing of their linguistic problems?
2. Do recasts and models help L2 learners notice their errors? If so, which feedback type is more effective?
3. Do modeling and recasting lead to any different learning effects? In other words, does the time between the utterance and the feedback make any difference in learning?
4. Do noticing the gap and noticing the hole lead to differential learning effects?

Two conditions (recasts and models) were set for this study. The term *model* in the present study refers to a sample of the target language provided after an interval of a few minutes, whose meanings reflect what the learner intended to convey. So, a model does not immediately follow the learner's ill-formed utterance; a recast does. In other words, a model is different from a recast in terms of its time of occurrence in relation to the learner's preceding utterance. The effects of recasts were compared with those of models to examine whether the elapsed time between the utterance and the feedback would make any difference. When learners notice holes in their production, both recasts and models will provide positive evidence to which they may readily pay attention, resulting in language learning. In this case, the difference between the model condition and the recast condition lies in how long learners can stay motivated to attend to the new input in order to receive their necessary information about the target language. When learners notice gaps through feedback, both recasts and models work as negative feedback. This provides negative evidence: information as to what is not allowed in the target language. In this case, the two conditions differ in that the learners in the model condition have to keep their original utterances in mind for a longer time than the learners in the recast condition do in order to make the cognitive comparison.

The model condition in this study was similar to the experimental condition in Izumi and Bigelow's (2000; Izumi et al., 1999) studies. The

learners in their experimental condition were involved in writing and then were provided with ready-made reading materials as the model input. The researchers pointed out that the degree of the discrepancy between the learner's production and the model input may influence the difficulty of noticing the gap (2000, p. 263). In contrast, for the present study, each learner in the model condition received the tailor-made model input, which made it easier to directly compare the output and the model input.

As described above, this study was designed taking the following points into account:

1. Two feedback types were provided: models and recasts. Recasts immediately followed the learners' utterances whereas models were provided a few minutes later.
2. To make models equivalent to recasts, models were created on the basis of the learners' utterances.

Method

Participants

Sixteen first-year university students (9 women and 7 men) volunteered to participate in this study. The participants were attending a required English class in a Japanese national university. They were Japanese learners of English as a foreign language (EFL); that is, all of them had the same L1 background (the Japanese language) and came from a similar learning environment (school subjects). The students were randomly assigned to one of the two groups: the model group ($n = 8$) and the recast group ($n = 8$). The model group included 5 women and 3 men (mean age = 18.4 years); the recast group included 4 women and 4 men (mean age = 18.5 years). The average years of English study was 6.06 for the model group and 6.75 for the recast group, ranging from 6 to 8 years; that is, most of the students had started to study English as a school subject when they were in the seventh grade, the first year of junior high school.

Procedure

Each student took part individually in the experimental session⁸ in June 2001. As can be seen in Table 1, the session consisted of three phases: Phase 1, Phase 2, and a retrospective interview. The total session

took approximately 30 minutes. The instruction and the retrospective interview were carried out in Japanese (L1).

Table 1. Experimental Sequence

	Model Group	Recast Group
	Explanation & Practice	Explanation & Practice
Phase 1	1. Picture Description Task - no feedback 2. Picture Identification Task	1. Picture Description Task - recasts/repetition
Phase 2	3. Picture Description Task	2. Picture Description Task
Phase 3	4. Retrospective Interview	3. Retrospective Interview

One by one, each student sat at a table across from me (a nonnative speaker of English whose L1 is Japanese). The session was videotaped and recorded with two tape recorders (one for the data collection and one for the interview). In Phase 1, the student was handed six cards. Each card had two similar pictures (Picture A and Picture B) on it (see Appendix). The student was told to select one of the two pictures and describe it in English so that I could identify that specific picture. Afterwards, the student was asked to move on to the next card and to repeat the procedure.

For the model group, I provided nodding along with expressions such as “I see,” “OK,” and “Yes.” In Example (2) from this study, Momoko tried to describe one of the pictures on Card 2 (Picture B). I demonstrated my understanding of her utterances by saying, “OK?” “Umm,” and “OK.” After she produced three utterances, I identified the picture (“Number 2-B”). Following the picture description task, the student in the model group performed the picture identification task. As in Example (3), I described the six pictures and the student identified those pictures. Through this task, model input was provided. During the previous picture description task, I took notes on what the student had said about the pictures and based the model input on these utterances in order to provide the model containing necessary information for the student. Therefore, the tailor-made model input was different for each student.

(2) Picture Description Task (the model group)

Momoko: *There are baby and woman.* (a)

Researcher: *OK?*

Momoko: *Baby ask ... what is that.* (b)

Researcher: *Umm.*

Momoko: *Woman say this is cat.* (c)

Researcher: *OK. Number 2-B.*

Momoko: *Yes.*

(3) Picture Identification Task (the model group)

Researcher: *There are a baby and a woman.* (a)

The baby asks what that is. (b)

The woman says that is a cat. (c)

Momoko: *Number 2-B.*

Researcher: *OK.*

The recast group performed only the picture description task. As in Example (4), I repeated every grammatical utterance (non-corrective repetition) or reformulated every erroneous utterance (recast). Both types of feedback were provided in the form of a confirmation check. Non-corrective repetition was provided to make the amount of feedback comparable to that of the model condition.

(4) Picture Description Task (the recast group)

Hikaru: *There is a boy and his mother.*

Researcher: *There are a boy and his mother? [recast]*

Hikaru: *Yes.*

Researcher: *OK.*

Hikaru: *The boy is reading a comic books on the sofa.*

Researcher: *Umm. The boy is reading a comic book on the sofa? [recast]*

- Hikaru: *Yes.*
- Researcher: *Umm.*
- Hikaru: *His mother angry.*
- Researcher: *His mother is angry?* [recast] *OK.*
- Hikaru: *Because he ... he doesn't study.*
- Researcher: *Because he doesn't study?* [non-corrective repetition]
- Hikaru: *Yes.*
- Researcher: *OK. Number 7-A.*
- Hikaru: *Yes.*

The description task was designed to elicit a variety of errors. Target structures were not determined for this study in order for the experimental task to be more natural. The task was considered to be communicative in that the student's main focus was primarily on the messages (focus on meaning) and because I did not know what the student was going to describe (information gap).

The two groups performed different tasks only in Phase 1. In Phase 2, the students were asked to describe the same six pictures that they had described in the picture description task of Phase 1. For example, because Momoko of Example (2) selected and described Picture B of Card 2 in Phase 1, she was asked to describe that picture (2-B) in English. In this second task, no feedback was provided. Because students described the same pictures in English in Phases 1 and 2, comparing the two phases was relatively easy. In the second trial, learning effects were examined by analyzing whether the students corrected the errors they had made in the first task. Then the retrospective interview followed.

The students were informed that the purpose of the study was to examine the processes of speaking. However, they were not provided with information about the sequence of the tasks in the session.

Measurement of Noticing

The operational definition of noticing used for this study was the "availability for verbal report" (Schmidt, 1990, p. 132; see also Leow, 1997).⁹ As Robinson (1995) pointed out, measuring noticing is not an

easy task because (a) “the experience of noticing may be fleeting and thus difficult to recall” and (b) “one may be aware of, yet unable to verbalize or otherwise articulate the nature of that which one is aware of” (p. 299). For this study, I needed to clarify whether L2 learners’ noticing would be brought about by feedback or not. In the stimulated recall (Gass & Mackey, 2000), learners are asked about what they were thinking at the time of feedback (Mackey et al., 2000). However, the stimulated recall did not seem to give sufficient information about noticing. For example, Mackey et al. (2000) provided an example of feedback that was perceived as phonological. The learner pronounced the word *flowers* as [flurs], and the interlocutor made the recast, “Floors?” In this case, the recast did not reformulate what the learner intended to say, and thus the learner had to reach the proper pronunciation by himself or herself. In the recall session, the learner stated, “I was thinking that my pronounce, pronunciation is very horrible” (p. 486). The researchers argued that this recall indicated that the feedback target was perceived as phonological by the learner. However, it is unclear from this recall protocol whether or not the feedback brought about the learner’s noticing of the wrong pronunciation. The learner might have already noticed his or her problem with pronunciation at the moment of the original utterance. For the present study, I decided to ask guided questions in the retrospective interview¹⁰ in order to clarify the noticing types.

As Gass and Mackey (2000) argued, during the retrospective interview, stimulus was provided to promote recall in order for the students to report verbally as much as possible. That is to say, during the retrospective interview, the tape-recorded interactions in the picture description task (Phase 1) were played back, with the pictures, in order to assist the students in verbally reporting their noticing about their linguistic problems or errors. The tape was stopped after each utterance, and the following questions were asked: (a) Are there any errors in your utterance?, (b) If so, when did you notice them?, and (c) What were you thinking at that time? When the students requested it, their recorded utterances were repeated to them.

Coding

Noticing Types

In this study, noticing was put into four categories, mainly on the basis of Swain’s (1998) noticing types: (a) problem, (b) production, (c) model/recast, and (d) a lack of noticing. The first type of noticing, coded

as *problem*, refers to noticing a hole in the interlanguage system. In most cases, the learners reported that in English they could not express what they wanted to say. In Example (5), Miho intended to say that the mother wanted the boy to study, but she had difficulty expressing this in English.¹¹ Thus, this verbal report indicated that Miho had noticed her linguistic problem at the moment of production.

(5) Noticing type 1: Problem

Tape: *(Her mother ... want ... he ... he mother want ... he ... he study ... homework.)*

Researcher: *Are there any errors?*

Miho: *Well, let's see, I don't know.*

Researcher: *At this moment, what ... what did you think?*

Miho: *Well, let' see, really.... As usual.... Like this. I wanted to say that this mother wanted this boy to study Somehow, with a that-clause... Like this ... as usual. What should I have said? I wanted to say she wanted this boy to study, but While I was saying in English, I found myself meaning that it was the mother who wanted to study Umm. I thought I should say it in a different way, but I didn't know how to express the meanings.*

In Example (6), Saori reported that what she had said in the task was wrong and that she had noticed the errors at the moment of production. When the learners noticed their errors at the moment of production, those reports were categorized as *production*. In the example, the learner did not report what she should have said in the target language (*old woman*). Some learners reported as follows, "I should have said *old woman*." This noticing type (production) is different from the first noticing type (problem) in that the learners did not state what they had intended to express; rather, they pointed out the errors in their utterances. This noticing type is identified with noticing the gap between the interlanguage and the target language without the help of feedback.

(6) Noticing type 2: Production

- Tape: (*There are old girl and boy.*)
- Saori: *There is something strange with the expres-
sion "old girl."*
- Researcher: *Yes. When did you notice it?*
- Saori: *When I was saying it.*

The third noticing type, *model/recast*, refers to noticing the gap with the help of feedback. In Example (7), Hikaru noticed the error in her utterance and reported that she had not inserted the "Be" verb *is*. Additionally, she stated that she had noticed the error through the feedback (recast for her).

(7) Noticing type 3: Model/recast

- Tape: (*His mother angry.*)
- Researcher: *What about this?*
- Hikaru: *I didn't say "is."*
- Researcher: *Yes. When did you notice that?*
- Hikaru: *When I heard you saying.*

Finally, those errors that the learners did not notice or the errors which they noticed at the moment of the interview were categorized as *a lack of noticing* (the fourth noticing type). In Example (8), Hikaru did not notice her error in using the noun phrase *a comic books*.

(8) Noticing type 4: A lack of noticing

- Tape: (*The boy is reading a comic books on the
sofa.*)
- Researcher: *What about this?*
- Hikaru: *Nothing particular.*
- Researcher: *No errors?*
- Hikaru: *No.*

Learning Effects

Learning effects were measured by analyzing the learners' performance in the second picture description task. When learners correctly produced utterances which had been problematic in the first task, those parts were categorized as *repair*. The repair category includes (a) the incorporation of the features provided in the model or recast and (b) the correct reformulation of the original problematic part, though the form is not identical with what the model or recast provided. The problematic parts, which still remained erroneous, were coded as *other*. This category contains (a) the same errors, (b) different errors, and (c) avoidance. Avoidance is referred to as non-occurrence of the linguistic context in question. These categories developed for this study were based on Lyster and Ranta's (1997) uptake categories.

Analysis

The learners produced a variety of errors, such as lexical errors, grammatical-morpheme errors, and syntactic errors (Sakai, 2002). The learners' pronunciation errors were not examined in this study because most of the learners' English was intelligible enough for me to understand probably because we shared the same L1 background and also because I did not intend to correct their pronunciation from the beginning.

The following analyses of the present study were based on *error points*, operationalized as the absolute number of errors identified by comparing the learner's utterance with the feedback.¹² The following example, Example (9), is part of Examples (2) and (3) described above. The model provided for the learner indicated several corrections of the learner utterance: (a) insertion of the definite article, (b) the 3rd-person present singular form, and (c) the word order in the subordinate clause. Thus, in the example, there were three error points.

(9) Error points (in the case of Momoko)

Learner Utterance: *Baby ask ... what is that.* [Three error points]

Model: *The baby asks what that is.*

I found verbal reports to vary among learners, even if they had received the same corrections through models or recasts. For example, some learners may make some comments only about the present 3rd-

person singular form whereas others may talk about all three corrections as in Example (9). In this study, therefore, the more detailed analyses were made on the basis of error points.

I coded all the data of the 16 learners. To assess inter-rater reliability for the coding, a subset (50%) of both the retrospective interview data and the performance data (Phase 2) was coded by another rater. For the coding of noticing types in the retrospective interview data, agreement reached 88.7% (258 out of 291); for the coding of learning effects in the performance data, agreement was 88.6% (272 out of 307).

Statistical Procedures

Because the sample size was small and the data were not normally distributed, I selected nonparametric procedures (Hatch & Lazaraton, 1991, p. 270). Medians were used as a measure of central tendency. Also, instead of standard deviations, this study used interquartile ranges (IQRs), which indicate the ranges of the middle 50% of the data. The Wilcoxon rank-sum tests were performed to test whether the two groups differed in terms of individual scores of error points and frequencies of feedback. For the analyses of noticing types and effects of learning, the chi-square tests were performed on the groups' data. When an expected cell frequency was less than five, Fisher's exact test was used (Hatch & Lazaraton, 1991, p. 409). Due to the small number of tokens of data, these analyses were carried out by merging individual data into two groups, model and recast.

Results

General Description

First of all, the two groups were compared in terms of error points in the picture description task of Phase 1 and the frequencies of feedback (models or recasts). Although this study did not control these factors in advance, the following general description data suggest that the learners in the two groups made errors and received feedback to a similar extent. Therefore, the two groups were considered to be comparable to each other in terms of the proficiency levels of learners and the provision of feedback.

Individual scores of error points in the picture description task of Phase 1 are summarized in Table 2. At first glance, the model group ($Mdn = 39.5$, $IQR = 9.5$) seemed to make more errors than the recast group did ($Mdn = 28.5$, $IQR = 15.25$). However, the Wilcoxon rank-sum

test showed that there was no statistically significant difference between the two groups ($p = .19$). Table 3 shows the numbers of models or recasts provided for each group. The results show that each student in both groups received about 20 of the models or recasts. The Wilcoxon rank-sum tests revealed no statistically significant differences between the groups ($p = 1.00$ for the numbers of models/recasts).

**Table 2. Error Points for Each Participant
(Picture Description Task of Phase 1)**

Model Group ($n = 8$)			Recast Group ($n = 8$)		
Participants	Sex	Error points	Participants	Sex	Error points
Yuzuko	female	51	Ayumi	female	43
Momoko	female	41	Saori	female	31
Makiko	female	41	Ryoko	female	26
Miho	female	38	Hikaru	female	17
Hanako	female	32	Akira	male	52
Taro	male	43	Noritake	male	38
Kyoji	male	32	Koji	male	24
Hideki	male	29	Ichiro	male	24

Note. All the names are fictitious.

Table 3. Number of Models/Recasts

	Model Group ($n = 8$)		Recast Group ($n = 8$)	
	<i>Mdn</i>	<i>IQR</i>	<i>Mdn</i>	<i>IQR</i>
Number of models/recasts	18	4.25	17	5.25

Analysis of Verbal Report on Noticing

The results of the analysis of verbal report are summarized in Table 4. The learners of the model group made a total of 307 errors. Of the errors, 26 were noticed as linguistic problems at the moment of pro-

duction (noticing the hole coded as problem), and accounted for 8.5%. Noticing the gap through models accounted for 3.9% (12 instances). On the other hand, a total of 255 errors were observed in the recast group. Noticing the hole coded as problem accounted for 7.5% (19 instances) while noticing the gap through recasts was 11.8% (30 instances). Both groups reported noticing holes in their interlanguage during production (8.5% and 7.5%): In other words, through production, the learners noticed what they had not been able to say in English. In both groups, most of the errors were not noticed by the learners (82.1% and 77.7%). One difference between the groups was that the recast group noticed errors through recasts more often than the model group noticed errors through models. A statistical analysis supported this.

The chi-square test revealed that there was a statistically significant difference in the distributions of noticing types between the two groups ($\chi^2(3) = 13.83, p < .01$). A further analysis of the residuals revealed that a significant difference existed only between the noticing of the gap through recasts by the recast group and the noticing of the gap through models by the model group ($p < .01$); no significant differences were found between the two groups in terms of problem, production, or a lack of noticing.

Table 4. Frequencies of Each Noticing Type

Noticing types	Model Group ($n = 8$)				Recast Group ($n = 8$)			
	Total	%	<i>Mdn</i>	<i>IQR</i>	Total	%	<i>Mdn</i>	<i>IQR</i>
Problem	26	8.5	2.0	4.25	19	7.5	2.5	3.00
Production	17	5.5	1.5	1.50	8	3.1	1.0	0.50
Model/recast	12	3.9	1.0	0.50	30	11.8	3.0	2.25
Lack of noticing	252	82.1	31.0	10.00	198	77.7	22.0	15.75
Total	307	100	39.5	9.50	255	100	28.5	15.25

Noticing Types and Learning Effects

Learning effects in relation to noticing types were measured as repairs in Phase 2. Table 5 indicates the repairs made by the model group. When the learners noticed their linguistic problems during their production in Phase 1, they repeated 57.7% of the errors (see the far left column under

the problem category). There were only two instances of incorporation from the model input. When they noticed their errors through models (see the middle column under the model category), they incorporated the target features from the input and repaired most of the errors (8 of the 12 errors, 66.7%). No instances were observed for the same error category. As for the production category, the results show a relatively high tendency to incorporate the information provided by the feedback (8 of the 17 errors, 47.1%).

Table 5. Learning Effects in the Model Group as Measured by Performance in Phase 2

Learning effects	Model Group				Total
	Problem	Production	Model	Lack of noticing	
Incorporation	2 (7.7%)	8 (47.1%)	8 (66.7%)	53 (21.0%)	71 (23.1%)
Self-repair	0 (0.0%)	1 (5.9%)	1 (8.3%)	7 (2.8%)	9 (2.9%)
Same error	15 (57.7%)	5 (29.4%)	0 (0.0%)	136 (54.0%)	156 (50.8%)
Different error	3 (11.5%)	3 (17.6%)	3 (25.0%)	27 (10.7%)	36 (11.7%)
Avoidance	6 (23.1%)	0 (0.0%)	0 (0.0%)	29 (11.5%)	35 (11.4%)
Total	26	17	12	252	307

The learning effects in the recast group are summarized in Table 6. When the learners noticed their linguistic problems by production, they produced the same errors consisting of 42.1% of the errors (see the far left column under the problem category). They incorporated the target features from the recasts, accounting for only 15.8% (three instances). When they noticed their errors through recasts (see the middle column of the recast category), 19 of the 30 errors were repaired (63.3%). They repeated four errors in Phase 2 (13.3%). The results of the Fisher's exact tests showed that there were no statistically significant differences in the distributions of repair in each noticing type between the two groups ($p = .68$ for problem; $p = .72$ for production; $p = .34$ for model/recast; $p = .88$ for lack of noticing).

Table 6. Learning Effects in the Recast Group as Measured by Performance in Phase 2

Learning effects	Recast Group				Total
	Problem	Production	Recast	Lack of noticing	
Incorporation	3 (15.8%)	4 (50.0%)	19 (63.3%)	36 (18.2%)	62 (24.3%)
Self-repair	1 (5.3%)	1 (12.5%)	0 (0.0%)	4 (2.0%)	6 (2.4%)
Same error	8 (42.1%)	3 (37.5%)	4 (13.3%)	115 (58.1%)	130 (51.0%)
Different error	3 (15.8%)	0 (0.0%)	6 (20.0%)	19 (9.6%)	28 (11.0%)
Avoidance	4 (21.1%)	0 (0.0%)	1 (3.3%)	24 (12.1%)	29 (11.4%)
Total	19	8	30	198	255

Because learning effects in each noticing type did not differ statistically between the two groups, I divided the noticing types into two groups as in Table 7 to clarify the differences of learning effects by noticing types. Noticing linguistic problems (the problem category) by the two groups was summed up in the left columns; noticing the gap through the models or recasts (the models/recasts category) by the two groups appeared in the right columns. Six of the 45 errors noticed as linguistic problems at the moment of production were repaired (13.3%); on the other hand, 28 of the 42 errors which had been noticed through the feedback were repaired (66.7%). The difference in repairs was statistically significant between the two noticing types ($\chi^2(1) = 25.95, p < .01$). In other words, noticing a gap through models or recasts led to more repairs in the later production than noticing a hole did.

Table 7. A Comparison of Learning Effects According to Noticing Types

Learning effects	Noticing the hole (problem)			Noticing the gap (recast/model)		
	Recast Group	Model Group	Total	Recast Group	Model Group	Total
Repair	4 (21.1%)	2 (7.7%)	6 (13.3%)	19 (63.3%)	9 (75.0%)	28 (66.7%)
Other	15 (78.9%)	24 (92.3%)	39 (86.7%)	11 (36.7%)	3 (25.0%)	14 (33.3%)
Total	19	26	45	30	12	42

Discussion

The results of this study are summarized as follows.

1. Production led the learners to notice their linguistic problems in the target language. Approximately 10% of the errors (8.5% for the model group; 7.5% for the recast group) were reported as linguistic problems at the moment of speaking (see Table 4).
2. Although both of the feedback types (models and recasts) helped the learners notice the gap between their interlanguage and the target language, recasts, which were provided immediately after their utterances, were more effective in leading to noticing than were models, which were provided a few minutes later, after their utterances (see Table 4).
3. There were no significant differences in learning effects by noticing types between the two groups (see Tables 5 & 6). Models and recasts may have worked in the same way in terms of learning effects.
4. Noticing the gap through feedback (models and recasts) showed a relatively higher effect in leading to repairs than noticing a hole did (see Table 7).

Roles for Output

The findings suggest that although L2 learners notice their linguistic problems by producing the target language, they might have much difficulty in attending to the subsequent input to take the necessary information in. According to the results (Table 4), the learners in this study reported that they had noticed their linguistic problems more frequently than the other noticing types, particularly in the model group. However, in terms of learning effects (Tables 5 & 6), noticing the holes led to no more repairs at the later production than a lack of noticing did. As in the previous studies (Izumi et al., 1999; Izumi & Bigelow, 2000), this study did not lend support to Swain's claim (1993, 1995, 1998) that through the activity of output, learners who notice their linguistic problems in their own interlanguage may have the need to pay closer attention to the subsequent input, resulting in learning. In this study, the input provided to the learners was exact reformulation of the learners' utterances, which

contained the structural information necessary for the learners. Therefore, even if provided with the necessary input, most of the learners did not utilize and incorporate the information in the model input provided in the form of models or recasts.

One might suppose that the learners did not retain what they noticed until the model input was given. Nevertheless, in fact, the time differential between noticing the hole and the subsequent input seems to have made no significant difference. In the recast condition, the input in the form of recast was provided immediately after the learners' utterances; in the model condition, the input in the form of a model was provided a few minutes later. Learning effects between the model condition and the recast condition were almost the same.

One possible explanation for the failure to make use of the subsequent input in later production may have been the developmental readiness of the learners (Mackey & Philp, 1998; Pienemann, 1998; Schmidt, 1990).¹³ In other words, the linguistic problems noticed by the learners may have been far beyond their current L2 competence. In Example (10), Saori stated that she had wanted to say, "The girl doesn't know whose watch that is." She came up with the question word *whose*, but did not know how to use the word in the sentence. That is, she faced a problem with subordination. As Pienemann (1998) pointed out, subordination is a construction in a relatively higher level of L2 development, which requires learners to process linguistic information beyond the clauses. It may be that because Saori's English abilities were not yet at the stage of subordination, she had difficulty in paying attention to and utilizing the subsequent input, even though she was aware of her linguistic problems.

(10) Noticing linguistic problems

Tape: *(She ... she ... um ... she don't ... don't ... she doesn't know ... um ... um ... she doesn't know watch ... whose watch...)*

Saori: *Well, I didn't come up with the expression of the girl who didn't know whose watch that was. Then, yes, "whose" occurred in my mind. I didn't know how to use "whose." A little bit irresponsibly, I did. Yes.*

Schmidt (1990) also points out that “availability for noticing and stages of L2 development are closely related” (pp. 142-143). The relationship between L2 learners’ readiness and ability to notice is beyond the scope of this study, but is worthy of further investigation.

Roles for Feedback

As to the relative effect of recasts and models on noticing, the results showed that noticing errors through recasts occurred more frequently than through models (Table 4). The difference between the two feedback types was the interval between the utterance and the feedback. As reported previously (Doughty & Varela, 1998; Long, Inagaki, & Ortega, 1998; Mackey & Philp, 1998), recasts make it easier for learners to compare their own utterances with the target structures. In other words, recasts are somewhat more effective in making cognitive comparisons than models are. Although models brought about noticing, the learners had more difficulty in making cognitive comparisons as time passed. This finding may support Long’s (1996) argument:

Negative feedback of this type (i.e., in the form of implicit correction immediately following an ungrammatical learner utterance) is potentially of special utility because it occurs at a moment in conversation when the NNS is likely to be attending to see if a message got across, and to assess its effect on the interlocutor. (p. 429)

As with learning effects, the model group and the recast group showed similar patterns in terms of repairing (see Tables 5 & 6). This may indicate that models functioned in the same way as recasts in terms of learning effects. The two feedback types are common in that they provide L2 learners with negative evidence and positive evidence at the same time. Probably this dual function of models and recasts results in similar learning effects. Also, it is possible that once learners notice errors through recasts or models, the provided feedback is likely to yield similar learning effects. Thus, Schmidt suggests that noticing plays an important role in mediating between the external information and the interlanguage system (Schmidt, 1990, 1993, 1995).

In summary, both recasts and models may similarly lead to repair in later production because of their common dual function of providing negative evidence and positive evidence simultaneously. Recasts, however, seem to enhance L2 learners’ noticing more effectively than

models do. In this study, feedback was provided artificially in a dense way, that is, every time the learners produced an utterance. This study suggests that feedback has the potential to make L2 learners notice the gap in interactions; nevertheless, the findings are not intended to be generalized to more natural situations.

Conclusions

In conclusion, Japanese EFL learners' verbal reports obtained through retrospective interviews were used to investigate whether they noticed their errors or linguistic problems during communicative tasks and, if so, when they noticed them. This study found that (a) although the learners noticed linguistic problems through the activity of speaking, they had difficulty in attending to and incorporating the subsequent input for the later production; (b) recasts were more effective in prompting noticing of errors than models were; and (c) noticing the gap through the feedback showed a relatively higher effect in leading to repairs than noticing the hole by output did.

Some limitations, however, need to be addressed. First, the proficiency levels of the participants were not available in this study. It cannot be assumed that the participants represented L2 learners of various proficiency levels. Therefore, I do not intend to generalize the findings of this study to all L2 learners, but rather I believe that this study provided some evidence on noticing by means of feedback or the activity of production. Secondly, the participants had been studying in my class at the time of the study. Our teacher-students relationship may have influenced their noticing. It is not clear, however, whether or not the relationship promoted the learners' noticing. Nevertheless, it must be pointed out that because in Japan most learners of English receive oral target language input primarily in the course of interactions in the classrooms, this experimental situation is not so unlike the learning environment of the participants.

Noticing can never be fully captured because it is the learner's internal process. However, if we hypothesize that noticing plays a mediating role between the input and the environment, then it becomes necessary to carry out research on the noticing issue and accumulate data in order to make a generalization. The findings of this study suggest that noticing the gap and noticing the holes may have differential effects on L2 learning.

Acknowledgments

An earlier version of this paper was presented at the 1st annual convention of the Japan Society of English Language Education (JASELE), Hiroshima, August 8, 2001. I would like to thank Professor Kinue Hirano of Joetsu University of Education, Ms. Carolyn L. Kaltenbach, as well as the editor and two anonymous reviewers of JALT Journal for their valuable suggestions and comments on an earlier draft of this paper. Any errors or omissions are mine.

Hideki Sakai is an associate professor at Shinshu University, Faculty of Education. His current research interests include noticing, feedback, and classroom second language acquisition.

Notes

1. Lyster (1998a) defined recasts as “the teacher’s implicit provision of a correct reformulation of all or part of a student’s ill-formed utterance” (p. 58). Thus, he seems to suppose that the role of interlocutor is limited to a teacher because his study focused on the classroom interactions. In contrast, Long (1996) stated that “*negotiation for meaning*, and especially negotiation work that triggers *interactional* adjustments by the NS [native speaker] or more competent interlocutor, facilitates acquisition.... Negotiation for meaning by definition involves denser than usual frequencies of semantically contingent speech of various kinds (i.e., utterances by a competent speaker, such as repetitions, extensions, reformulations, rephrasings, expansions and recasts), which immediately follow learner utterances and maintain reference to their meaning....” (pp. 451-452). Thus, he seems to consider that recasts are not necessarily provided by a native speaker. Although an anonymous reviewer pointed out the possibility that L2 learners provide recasts to each other, it is not clear whether L2 learners of the same proficiency level can perform recasts of each other’s utterances. This would require empirical study in the future.
2. It should be noted that the definition of recasts varies in studies on the effects of recasts: As Nicholas, Lightbown, and Spada (2001) stated, “one important difference in operational definitions in L2 studies is that recasts have sometimes included more than just a reformulation of a learner’s incorrect utterance” (p. 749). For example,

in Doughty and Varela's (1998) study, recasts were accompanied by repetition of the learner's ill-formed utterance. Doughty and Varela (1998) termed this combination of feedback as "corrective recasting" (p. 123). Mackey and Philp (1998) compared the two conditions between interaction with intensive recasts and interaction without intensive recasts. The former condition was described as "the artificial 'flooding' of interaction with recasts" (p. 353), resulting in 350% more than the latter condition. Long, Inagaki, and Ortega (1998) seemed to follow the definition of recasts as those without any emphasis, but their results were ambiguous. They carried out two studies and found the superiority of recasts over the provision of samples of the target structures, which they called modeling, in only one of the two structures in one of the studies. At present, therefore, it is quite difficult to argue for the effects of recasts alone based on these studies. I am grateful to an anonymous reviewer for pointing this out to me.

3. The learner names in this paper are fictitious.
4. According to Lyster (1998a), confirmation checks are those parts of recasts, which reformulate "all or part of the utterance with rising intonation and no additional meaning" (p. 48). As stated in the Procedure section, recasts were provided with rising intonation in this study, that is, equivalent to confirmation checks.
5. Lyster and Ranta (1997) referred to the learner's response to the feedback provided by the teacher as "uptake" (p. 49) and found that the percentage of uptake of recasts was 31%, which was quite low compared to the other types of feedback: elicitation, clarification request, metalinguistic feedback, explicit correction, and repetition. More relevant to the argument in the present study is Lyster's (1998a) finding that 95% of the non-corrective repetition did not lead to uptake (p. 67). This provides another piece of empirical evidence suggesting that L2 learners respond to recasts and non-corrective repetition differently.
6. Gass and Mackey (2000) considered stimulated recall to be "one subset of a range of introspective methods that represent a means of eliciting data about thought processes involved in carrying out a task or activity" (p. 1). In other words, introspection ranges from concurrent to retrospective think-aloud protocols.

7. I am using the term models in a narrower sense than in Long, Inagaki, and Ortega's (1998) study that compared the effects of recasts and models. In their study, models were pre-determined recorded samples of the target structure and functioned as only positive evidence. As a result, the propositional content of the models was out of the learners' control. In the current study, models were created on the basis of the learners' utterances.
8. Prior to the start of data collection, pilot sessions were carried out with three adult learners in order to develop and refine the experiment procedure. These learners were not included in the analysis of the current study.
9. There have been some controversies as to whether noticing entails awareness or not (Robinson, 1995; Simard & Wong, 2001; Tomlin & Villa, 1994). Tomlin and Villa (1994) argued that attention consists of three separable functions: alertness, orientation, and detection (p. 198). Among the three functions, they considered detection, defined as "the cognitive registration of sensory stimuli" (p. 192), to be the most significant because detected information can be processed further for learning. Moreover, they claimed that "such detection does not require awareness" (p. 199). Robinson (1995), based on memory studies, stated that "noticing can be identified with what is both detected and then further activated following the allocation of attentional resources from a central executive" (p. 297). Schmidt (1993) himself considered noticing to be "related to rehearsal within working memory and the transfer of information to long-term memory, to intake, and to item learning" (p. 213). Although he admitted that detection does not require awareness, Robinson (1995) pointed out that "subliminal exposure effects are unlikely to have effects over intervals longer than a few hundred milliseconds, are rapidly lost from memory, and cannot in any useful sense be claimed to be evidence of learning" (p. 298). In addition, Simard and Wong (2001) reviewed the studies cited by Tomlin and Villa (1994) and called into question Tomlin and Villa's (1994) position that awareness is dissociated from detection. At the moment, it can be agreed that noticing, operationally defined as "availability for verbal report" (Schmidt, 1990, p. 132), is necessary for further processing of learning, that is, SLA.
10. I basically followed the stimulated recall methodology guidelines described in Gass and Mackey (2000). However, I refer to the

recalling session in this study as *the retrospective interview* because guided questions were provided. It should be noted that the guided questions may have brought about validity problems in the recalling procedure, that is, nonveridicality (Gass & Mackey, 2000, p. 107). These problems should be kept in mind while interpreting the results.

11. The participants spoke in Japanese at the interview sessions. In this paper I gave examples of their protocols translated into English. Analysis was carried out on the original protocols.
12. An anonymous reviewer pointed to the problems of identifying errors and stated that “We are told that ‘Baby ask ... what is that’ [as in Example (9)] contains three errors. But ‘what is that’ is perfectly correct if the learner is using direct speech and we have no way of knowing whether the learner intended direct or reported speech.” Indeed it is extremely difficult to identify errors. For this study, I used an external criterion; that is, changes made in the feedback were counted as errors. Thus, the identification of errors was performed objectively in this study. In the case of Example (9), the recast reformulated the learner’s utterance in three ways. Based on the changes made in the recast, the error points were identified. So I did not calculate the inter-rater reliability for the identification of errors. This method of identification of errors may be considered to be valid because the study attempted in part to examine how the learners utilize the information provided through feedback.
13. As an anonymous reviewer pointed out, another possible explanation is that L2 learners might have had difficulty in attending to form and meaning simultaneously (VanPatten, 1990). However, it should be noted that in the second picture description task, the students described the pictures that they had already worked on in the first picture description task. Thus, the second task was not new to L2 learners in a true sense. This probably lessened cognitive load in the second task.

References

- de Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language Learning*, 46, 529-555.
- Doughty, C. (1994). Fine-tuning of feedback by competent speakers to language learners. In J. Alatis (Ed.), *GURT 1993: Strategic interaction and language acquisition* (pp. 96-108). Washington, DC: Georgetown University Press.

- Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114-138). Cambridge: Cambridge University Press.
- Doughty, C., & Williams, J. (1998). Pedagogical choices in focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 197-261). Cambridge: Cambridge University Press.
- Ellis, R. (1994). A theory of instructed second language acquisition. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 79-114). San Diego, CA: Academic Press.
- Gass, S. M. (1988). Integrating research areas: A framework for second language studies. *Applied Linguistics*, 9, 198-217.
- Gass, S. M. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. New York: Newbury House.
- Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *TESOL Quarterly*, 34, 239-278.
- Izumi, S., & Bigelow, M. (2001). Methodological and theoretical issues in testing the effects of focus on form. *TESOL Quarterly*, 35, 181-189.
- Izumi, S., Bigelow, M., Fujiwara, M., & Fearnow, S. (1999). Testing the output hypothesis: Effects of output on noticing and second language acquisition. *Studies in Second Language Acquisition*, 21, 421-452.
- Leow, R. P. (1997). Attention, awareness, and foreign language behavior. *Language Learning*, 47, 467-505.
- Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. B. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39-52). Amsterdam: John Benjamins.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Richie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). New York: Academic Press.
- Long, M. H., Inagaki, S., & Ortega, L. (1998). The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *Modern Language Journal*, 82, 357-371.
- Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15-41). Cambridge: Cambridge University Press.
- Lyster, R. (1998a). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20, 51-81.
- Lyster, R. (1998b). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion programs. *Language Learning*, 48, 183-218.

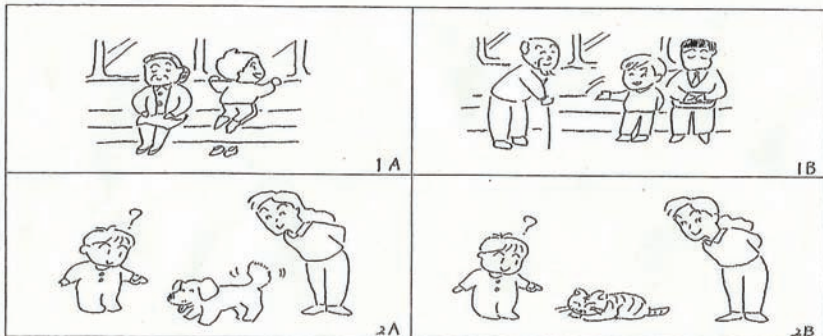
- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19, 37-66.
- Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, 22, 471-497.
- Mackey, A., & Hilp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *Modern Language Journal*, 82, 338-356.
- Nakamura, A. (1995). *CD tsuki jidoeiken kuria mondaishu*. Tokyo: Seibido.
- Nicholas, H., Lightbown, P. M., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning*, 51, 719-758.
- Oliver, R. (1995). Negative feedback in child NS/NNS conversations. *Studies in Second Language Acquisition*, 17, 459-481.
- Pienemann, M. (1998). *Language processing and second language development: Processability theory*. Amsterdam: John Benjamins.
- Roberts, M. A. (1995). Awareness and the efficacy of error correction. In R. Schmidt (Ed.), *Attention & awareness in foreign language learning* (pp. 163-182). Honolulu: University of Hawaii at Manoa.
- Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning*, 45, 283-331.
- Sakai, H. (2002). Error types and noticing: Roles for feedback and L2 learner repair. *Bulletin of Joetsu University of Education*, 21, 741-755.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129-158.
- Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, 13, 206-226.
- Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1-63). Honolulu: University of Hawaii at Manoa.
- Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237-326). Rowley, MA: Newbury House.
- Simard, D., & Wong, W. (2001). Alertness, orientation, and detection: The conceptualization of attentional functions in SLA. *Studies in Second Language Acquisition*, 23, 103-124.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.

- Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review*, 51, 158-164.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honor of H. G. Widdowson* (pp. 125-144). Oxford: Oxford University Press.
- Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64-81). Cambridge: Cambridge University Press.
- Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal*, 82, 320-337.
- Tomlin, R., & Villa, H. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, 16, 183-203.
- Truscott, J. (1998). Noticing in second language acquisition: A critical review. *Second Language Research*, 14, 103-135.
- Van den Branden, K. (1997). Effects of negotiation on language learners' output. *Language Learning*, 47, 589-636.
- VanPatten, B. (1990). Attending to form and content in the input: An experiment in consciousness. *Studies in Second Language Acquisition*, 12, 287-301.
- Whitlow, J. (2001). Comments on Shinichi Izumi and Martha Bigelow's "Does output promote noticing in second language acquisition?": Some methodological and theoretical considerations. *TESOL Quarterly*, 35, 177-181.

(Received July 3, 2002; revised May 31, 2003)

Appendix

Sample Picture Cards Used (Revised from Nakamura, 1995)



Using Mnemonics to Facilitate Learning of Japanese Script Characters

Emmanuel Manalo

The Student Learning Centre, The University of Auckland

Satomi Mizutani

*Department of Applied Language Studies and Linguistics,
The University of Auckland*

Julie Trafford

The Student Learning Centre, The University of Auckland

Extra tutorial sessions on the use of pictorial mnemonics to facilitate memorization of Japanese hiragana and katakana script characters, as well as vocabulary words and kanji characters, were offered to university first year undergraduate students taking a beginners' Japanese language course. 27 students, most of whom were experiencing some difficulties with the course, volunteered to attend. Although the improvement in actual marks that the students evidenced subsequent to attending the sessions did not prove to be statistically significant, a significant improvement in pass rate was found. Furthermore, the students rated the sessions highly in terms of their helpfulness, and the majority indicated that they believed the sessions helped their performance in the course assessments. It is concluded that mnemonic strategies can effectively be employed in facilitating retention of the script of a foreign, non-alphabetic language within a real educational setting.

初心者対象の日本語コースを取っている大学1年生に、日本語のひらがなとカタカナ、及び語彙と漢字の記憶を促進するため、絵を用いた連想法を使った追加授業が行われた。27名（そのうちのほとんどは、コースにおいて何らかの困難に直面している）が自主的に追加授業に出席した。追加授業出席後、学生の実際的小テストの点の向上には有意差は認められなかったものの、合格率の向上においては有意差は確認された。さらに、追加授業に出席した学生は授業が役立ったと高く評価し、大多数が追加授業がコースの成績向上に貢献したと思うと述べた。本論は、実際の教育現場で英語のアルファベットを用いない外国語の文字を教える際、連想法を効果的に使用できると結論づけた。

Mnemonics are schemes for assisting memory. They include well-known and simple rhymes like the one that starts “Thirty days hath September, April, June and November ...,” intended as a reminder of the number of days in each of the months of the year, as well as methods that are a little more technical such as *soh cah toa*, a code that for many students of trigonometry serves as a reminder of when to use the sine, cosine, and tangent rules for solving the dimensions of triangles. Baddeley (1997) provided a good general description of mnemonic strategies, and noted that they usually employ various manipulations of the material to be remembered to make that material more memorable (e.g., by using imagery, elaboration, or reduction).

Research findings during the past twenty years have shown that mnemonic strategies can have practical educational applications. Their usefulness in learning a wide range of information, from foreign vocabulary words to science facts, have been reported in many research studies (e.g., Desrochers, Gelinas, & Wieland, 1989; Ehri, Deffner, & Wilce, 1984; McDaniel & Pressley, 1989; McDaniel, Pressley, & Dunay, 1987; Rosenheck, Levin, & Levin, 1989). These studies consistently show that participants taught using mnemonic methods outperform those taught using other methods in tests that gauge the retention of the target information (for reviews, see Levin, 1983; and Manalo, 1997). Furthermore, mnemonics have been found effective in teaching individuals with learning disabilities (e.g., Conodus, Marshall, & Miller, 1986; Elliott & Gentile, 1986; Manalo, 1991; Manalo, Bunnell, & Stillman, 2000; Mastropieri, Scruggs, & Fulk, 1990; Mastropieri, Scruggs, Levin, Gaffney, & McLoone, 1985), as well as those who have sustained brain injury (for a review, see Richardson, Cermak, Blackford, & O’Connor, 1987). Thus mnemonics can provide a potentially effective alternative instruction strategy when other more traditional forms of instruction may have already been tried but failed to produce the desired results in the learning performance of individuals who have special needs. A number of recent research studies have also shown that, contrary to traditional views that mnemonics are unnatural and used only in artificial environments, the majority of people in general spontaneously use various forms of mnemonics when given tasks that require memorisation (e.g., Brooks, Friedman, Gibson, & Yesavage, 1993; Hill, Schwob, & Ottman, 1993; Manalo, 1999).

Language learning is an area of educational pursuit that, among other things, requires a considerable amount of memorisation. It is therefore not surprising that there are many mnemonic techniques advocated and practised by instructors and students alike in this area. Research, how-

ever, has focused mainly on vocabulary acquisition and the use of the keyword method. Atkinson and Raugh (1975) provided a good description of the keyword method, as well as an application of it in learning Russian vocabulary. As they explained, using this method to learn a new word involves two basic steps: finding an English word (the “keyword”) that sounds similar to a part of the foreign vocabulary word to be learnt, and then creating a mental image of the keyword interacting with the English translation of the foreign word. Hence, to learn, for example that the Russian word *zdanie* (pronounced “z-dawn-yeh”) means *building* in English, the English word *dawn* is used as a keyword, and then a mental image of “the pink light of dawn reflected in the windows of a tall building” (Atkinson & Raugh, 1975, p. 127) is conjured. Thus, when the word *zdanie* is encountered again in the future, the image previously conjured would return, serving as a reminder of the correct translation of the word in English. The keyword method of course is not limited to English translations: the user’s native or other familiar language can be used in generating keywords for association with the translation of the foreign or other unfamiliar word to be learnt.

A good example of a study that examined the effectiveness of the keyword method was reported by Levin, McCormick, Miller, Berry, and Pressley (1982). In this study, the participants—children in the fourth grade—were required to learn the meaning of unfamiliar words, such as that *surplus* means “having some left over, having more than was needed.” Children in the mnemonic condition were shown pictures that involved stimulus recoding. For example, they were shown a picture of someone pouring lots of syrup over pancakes and saying that there was a surplus of it in the cupboard. Thus, apart from illustrating a situation where surplus was involved, the word *syrup* was also used as a keyword to help remember the meaning of the word surplus. Levin et al. found the mean vocabulary recall of children in the mnemonic condition significantly greater than the recall of children who were shown non-mnemonic pictures, and children who were shown no pictures at all (i.e., just the words and their definitions, and either verbal contexts or nothing else).

Other research studies have found mnemonic keywords effective in facilitating the acquisition of German nouns and their grammatical gender (Desrochers, Gelinis, & Wieland, 1989), the learning of the meanings of unfamiliar Old English words (McDaniel & Pressley, 1989) and obscure English words (McDaniel, Pressley, & Dunay, 1987), and the recall of botany concepts (Rosenheck, Levin, & Levin, 1989). In fact, in reviewing the effectiveness of the keyword method in vocabulary learn-

ing, Levin and Pressley (1985, p. 153) went as far as stating that “with respect to definition memory, no strategies investigated to date ... have rivaled the mnemonic keyword method in either their consistency or their potency.”

Whilst there have been many studies that have investigated the effectiveness of mnemonics in vocabulary learning, there have only been three previous research studies that have looked at the usefulness of mnemonics in learning the script of another language: Gruneberg and Sykes (1996), Lu, Webb, Krus, and Fox (1999), and Quackenbush, Chujo, Nagamoto, and Tawata (1989). Gruneberg and Sykes used mnemonic descriptions aimed at helping learn letters of the Russian alphabet and their equivalent English sounds. For example, the Russian letter ϕ sounds like the English letter F, and so experimental participants were asked to imagine ϕ as looking like a fishing float. Gruneberg and Sykes found that experimental participants obtained a significantly higher overall score compared to control participants (who were not provided with the mnemonic descriptions) in subsequent tests that required them to recall the English equivalents of the Russian letters. Lu et al.’s results were similar: In their case, they found that their participants learned more of Japanese and Chinese kanji characters and their meanings when the characters were presented with the aid of descriptive mnemonics. For example, they used the description “Three peaks of a MOUNTAIN” as a mnemonic to help learn the kanji character for mountain, 山.

While both the Gruneberg and Sykes (1996) and the Lu et al. (1999) studies reported better memory performance from participants provided with mnemonic descriptions, both studies used artificial (laboratory) rather than real classroom settings, and their participants were not really learning the languages in question. Both studies also did not require participants to produce or write the scripts in question, as students who are really studying these languages would usually be required to do. Thus, as Gruneberg and Sykes themselves acknowledged, the usefulness of their findings “in real life learning situations is difficult to assess” (p. 83). There are in fact very few studies that have looked at the usefulness of mnemonic strategies where actual production of the foreign language is involved. Ellis and Beaton (1993) is one such study, and their findings suggest that—at least where the keyword method is concerned—a mnemonic strategy may not be as effective in facilitating recall of the foreign word (given the native translation of the word as stimulus, e.g., English to German), as it appears to be in facilitating recall of the native translation of the foreign word (e.g., German to English).

It should also be noted that the Gruneberg and Sykes (1996) and Lu et al. (1999) studies did not examine the potential usefulness of *pictorial* mnemonics in learning the script of another language. Although the mnemonics employed in both studies largely pertained to the visual characteristics of the scripts in question, both employed *descriptive* rather than pictorial mnemonics. It could of course be argued that the descriptive mnemonics provided to the experimental participants lent themselves to visualization. However, the extent to which participants employed visualization was not gauged in either of the studies. Levin (1983) argued that a beneficial feature of pictorial mnemonics is that “the relationship between each letter’s visual and phonetic properties is strengthened by an easily identified picture that contains an analogous relationship” (p. 219). He gave examples of the letters M, F, and S transformed to look like a mountain, a flag, and a snake respectively in their pictorial mnemonic states. Whether this beneficial feature of pictorial mnemonics holds true for the learning of script characters other than English was not investigated in the Gruneberg and Sykes, and Lu et al. studies. In any case, while Levin and Pressley (1985) acknowledged that although the pictorial component is not necessary for mnemonic strategies to be effective, they believed that the use of pictures reduces “the information-processing load associated with visual imagery generation” (p. 158).

It needs to be noted that the use of pictorial mnemonics for remembering the script of another language is advocated in many books and by some language teachers. In teaching Japanese alone, there are teachers’ guides such as Ogawa’s (1990) *Kana can be easy*; Quackenbush and Ohso’s (1983) *Hiragana in 48 minutes*; Quackenbush and Ikeda’s (1989) *Katakana in 48 minutes*; Quackenbush’s (1999) *Hiragana/katakana in 48 minutes: Teacher guide*; and Takabe’s (1993) *Kanji isn’t that hard! Kanji can be mastered with the 24 rules*. These types of instruction guides are used by many language instructors despite the fact that hardly any research has been undertaken and published to report on the efficacy of such techniques. The authors of the present paper are aware of only Quackenbush, Chujo, Nagamoto, and Tawata’s (1989) paper reporting on a study that attempted to find out if pictorial mnemonics are effective in learning one of the Japanese scripts, hiragana.

Quackenbush et al. (1989) used 16 foreign students with little knowledge of the Japanese language, who were enrolled in a Japanese language course at Hiroshima University, as participants for their study. They assigned nine of the students to a “mnemonics” group (taught the hiragana script using Quackenbush and Ohso’s book), and seven to a “colour

association" group (taught using a colour matrix so that each hiragana character had a top line colour and a right hand side colour derived from the matrix). Quackenbush et al. found that the mnemonic and colour association groups did not significantly differ in their pre-instruction scores in hiragana *reading*. Furthermore, even though both groups significantly improved in their post-instruction scores in hiragana reading, no significant advantage was shown by the mnemonics group. After the post-instruction test, Quackenbush et al. allowed the participants three days of study at home before administering a test in hiragana *listening*. Here they found the mnemonic group scored significantly better compared to the colour association group. However, because no pre-instruction scores were collected and reported on the groups' *pre-instruction* hiragana *listening* performance, their conclusion that mnemonic instruction facilitated better long term retention is unwarranted.

There are a number of other problems with the Quackenbush et al. (1989) study that make it impossible to gauge the usefulness and efficacy of pictorial mnemonic instruction in learning the Japanese hiragana script. One of the more major of these problems lies with the questionable rationale for providing colour association instruction to the other group instead of using a control group provided with standard classroom instruction in the hiragana script. Quackenbush et al. provided no research evidence to show the colour association strategy to be more effective than (or at least as effective as) standard classroom instruction. There is therefore no possible way of telling whether the pictorial mnemonic instruction they used improved the students' acquisition of the Japanese hiragana script compared to what would normally have been expected.

Thus, despite the three studies that have investigated the usefulness of mnemonic strategies in learning the script of another non-alphabetic language, a number of important questions remain unanswered. First, are such strategies useful in real classroom settings, with students who are really learning the language in question and who not only have to recognize but also recall and reproduce the script? Second, would evidence of the usefulness of mnemonic strategies manifest in class assessment scores and in student appraisal of those strategies? Third, are pictorial mnemonics—in particular—useful in remembering the script of another language, as is advocated in many books and by some language teachers?

The present study sought to address the above questions where the learning of Japanese script characters in a natural university classroom setting was concerned. Thus the main hypothesis was that students tak-

ing a first year Japanese course who attend extra tutorials showing them how pictorial mnemonics could be used to remember Japanese script characters would evidence an improvement in their course assessment.

Method

Participants

The participants in this study were 27 students who were taking a stage one Japanese language acquisition course at a university in New Zealand. The course is designed for students with little or no background in learning the Japanese language and includes instruction and assessment in reading, writing, and speaking. It comprises four hours of instruction per week over 12 teaching weeks. The 27 students volunteered to attend extra tutorials that were advertised as being aimed at those who may be experiencing some difficulties in learning the Japanese script. Seventeen of the 27 students were female, and 10 were male. Their mean age was 20.8 years ($SD = 2.95$), with the range being 17 to 28 years. The other 169 students who were taking the course but did not attend the extra tutorials served as a control group.

The students involved in this study (participants and controls) came from a variety of first language (L1) backgrounds, reflecting the diverse multicultural mix of students at the university where this study took place. However, all were proficient in English, which is a requirement for entry at the university.

There are a number of guidelines under which a study of this kind—in natural settings—needs to operate if fairness to the participants (and potential participants) is to be considered and the study is to be approved by the appropriate institutional ethics committee. For example, in the present study, because the participants were students who were really studying the language in question, random assignment to an experimental group was out of the question: the participants had to volunteer and hence self-select to be in the experimental group (exclusion from attending potentially beneficial instruction is not acceptable practice). Likewise, providing delayed or placebo instruction was not permissible as this could conceivably disadvantage students in their course performance. The control group in this study is therefore merely a comparison group for establishing the stability of the students' course performance in the absence of the experimental instruction provided (see Brown, 1992, for the use of this kind of control group in language research). Whilst there are many obvious restrictions and limitations about how

variables affecting the participants can be manipulated or controlled in a study like this (in a natural setting), there are also numerous advantages which include having participants with the “real” motivations for what they are learning, and greater confidence about the generalisability of findings. Apart from research in language teaching and learning, academic intervention studies also fall under this category (e.g., Heerman & Maleki, 1994; Manalo, Wong-Toi, & Henning, 1996; Walsh, 1985).

Procedure

In the Japanese course that the students were enrolled in, there were regular course assessments in the form of short tests administered during regular class times. There were six such tests (one each week) during the first half of the semester, when the present study was undertaken. These tests covered vocabulary, grammar, and kanji, and were usually held on a Tuesday or Wednesday. Tests 1 and 2 did not require writing in Japanese script. From Test 3 onwards, writing in Japanese script was required. The extra tutorials covering the use of pictorial mnemonics were offered after Test 3. The timing of the extra tutorial sessions in relation to the course weeks and tests, and the content of these tutorials, are shown in Table 1. At least two streams of each extra tutorial session were offered to avoid possible clashes with the students’ regular lectures and tutorials.

Pictorial mnemonics that the students could use to better remember the Japanese script characters, in hiragana and katakana, were described, drawn on the board, and explained during the extra tutorial sessions. For example, the hiragana symbol *い* is approximately equivalent to the short /i:/ sound of the vowel *i* and was described as “two eels” (see Figure 1). The pictorial mnemonics used were a combination of those devised by Ogawa (1990), Quackenbush (1999), and ones devised by the second author of this paper. In the examples provided in Figure 1, the pictorial mnemonics for *い* is based on Ogawa’s pictures, *い* is based on Quackenbush’s pictures, and the mnemonic for *い* was devised by the second author.

During the tutorial session on using mnemonics to remember vocabulary words (extra tutorial session number 6), the keyword method—where new foreign words to be learned are associated with similar sounding words known to the learner—was described to the students. Ten examples, such as those shown in Figure 2 (which were devised by the first author), were explained and drawn on the board. It was pointed out to the students that from the key words extracted (e.g., “sago” and

Table 1. Schedule of Course Tests and the Extra Tutorial Sessions Provided on Mnemonic Use (with Weeks of the Course Indicated in Brackets)

Tests 1 to 3 (Weeks 1-3)	
Extra tutorial sessions (Weeks 3-4):	1 Hiragana あ (A) to の (NO) [25 Japanese script characters]
	2 Hiragana は (HA) to ん (N) [21 Japanese script characters]
	3 Revision of hiragana and variations
Test 4 (Week 4)	
Extra tutorial sessions (Weeks 4-5):	4 Katakana ア (A) to ノ (NO) [25 Japanese script characters]
	5 Katakana ハ (HA) to ン (N) [21 Japanese script characters]
Test 5 (Week 5)	
Extra tutorial sessions (Weeks 5-6):	6 Vocabulary
	7 Kanji
Test 6 (Week 6)	

“hand” from *asagohan*), mnemonic images and/or phrases could be constructed to help in remembering the target foreign word. The importance of practising recall and self-testing was also emphasised. As an exercise, the students were given a list of 12 new Japanese vocabulary words to “extract” key words from and then construct mnemonic mental images and/or phrases for. They were then subsequently asked to recall the Japanese words when prompted with the English equivalents.

During the session on mnemonics for remembering kanji (extra tutorial session number 7), the students were introduced to methods with which they could better appreciate the meaningful nature of kanji script characters and construct pictorial mnemonics, when necessary, to help in remembering these script characters. Approximately 30 examples (including variants), such as those shown in Figure 3 (based on mnemonics devised by Takabe, 1993, and by the second author), were explained, drawn on the board, and discussed. The students were also given an opportunity to practise applying the strategies described to about 20 kanji script characters.

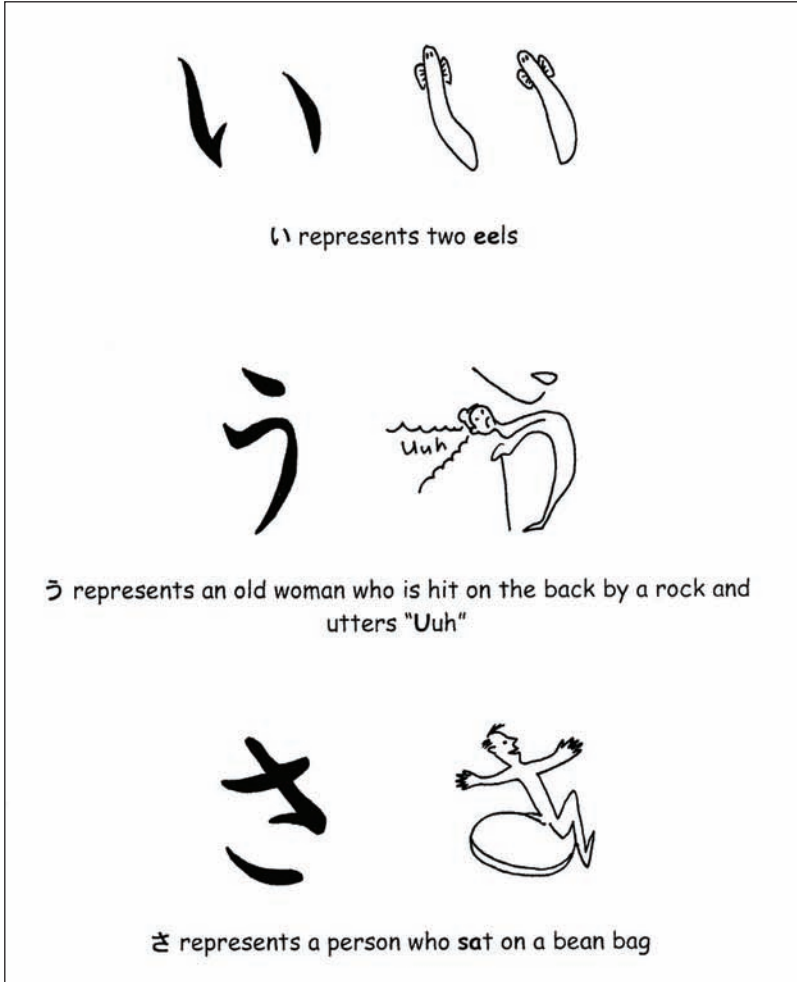


Figure 1. Examples of pictorial mnemonics used for the hiragana script characters

The students were able to attend the extra tutorial sessions at any stage during the time they were offered: 16 of the 27 participants started attending prior to Test 4, eight participants started attending prior to Test 5, and three attended for the first time before Test 6. The participants' pre- and post-attendance test scores and average pass rates were examined. Subsequent to the extra tutorial sessions, the participants were also

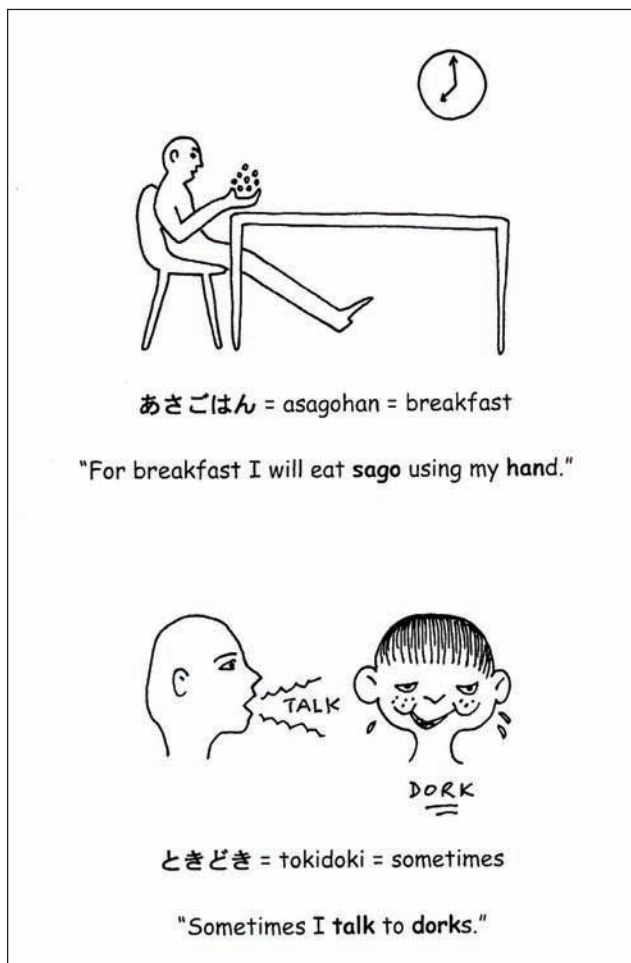


Figure 2. Examples of mnemonics used for the Japanese vocabulary words

asked to complete a questionnaire evaluating the usefulness of the sessions provided, and the Academic Motivation Scale (Vallerand, Pelletier, Blais, Briere, Senecal, & Vallieres, 1992).

The authors are well aware of the debate concerning the value of student evaluations of courses they take. However, while some authors are very negative about the value of such evaluations, citing reasons

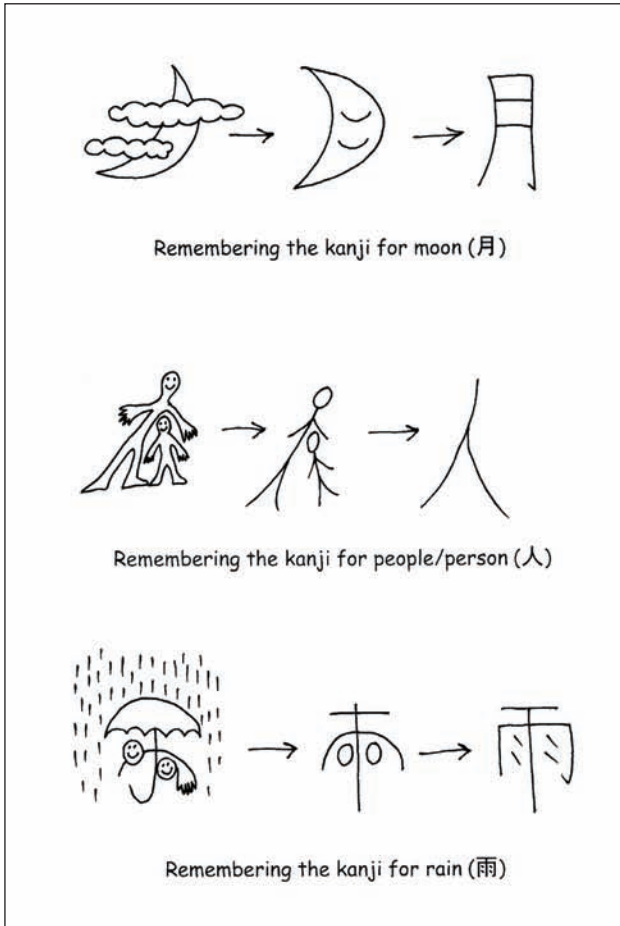


Figure 3. Examples of mnemonics used for the kanji script characters

such as instructors bowing to student demands in order to gain satisfactory course evaluations for the purposes of tenure or promotion (e.g., Trout, 1997), the students' honesty hardly ever comes into question. It is generally accepted that, at least where *anonymous* course evaluations are concerned, the students filling them out will be honest—in fact, sometimes brutally honest, particularly when they are not satisfied with a course. As DeZure (reported by Plater, Matthews, D'Appollonia,

& Abrami, 1997, p. 17) pointed out, “students want opportunities for mutual ongoing feedback, including a chance to tell their instructors what is working for them, and what is not.” Using clustering procedures, Young and Shaw (1999) also found that students’ perception of the value of a course is one of the three most important variables that impact on their evaluation of it.

The Academic Motivation Scale was administered to find out whether the participants differed in any way in their academic motivation from their counterparts who chose not to attend the extra tutorial sessions. Hence, amongst the “control” group of students who did not attend the extra tutorials, volunteers were solicited to likewise complete the Scale, and 48 students obliged. The Scale provides scores in intrinsic motivation (engaging in an activity for itself, and the pleasure and satisfaction that could be derived from undertaking that activity), extrinsic motivation (engaging in various kinds of behaviour as a means to an end and not for their own sake), and amotivation (when no contingencies are perceived between outcomes and one’s own actions) (Vallerand et al., 1992).

The students in both groups were not required to write their names on either the questionnaire to evaluate the usefulness of the tutorials or the Academic Motivation Scale.

Results

As noted in the previous section, students were able to attend the extra tutorial sessions offered at any stage. The participants’ test scores were therefore analysed according to whether they were pre- or post-attendance scores. For example, a participant who started attending from extra tutorial session 1 would have his/her scores in Tests 1 to 3 categorised as pre-attendance, while his/her scores in Tests 4 to 6 would be categorised as post-attendance. On the other hand, a participant who started attending only from extra tutorial session 6 would have his/her scores in Tests 1 to 5 categorised as pre-attendance, and only his/her score in Test 6 categorised as post-attendance. (Please refer to Table 1 for clarification on when the tests occurred in relation to the extra tutorial sessions.)

Test performance

The participants’ mean pre-attendance test score was 51.39% ($SD = 28.53$), while their mean post-attendance score was 55.81% ($SD = 26.74$).

Although these scores suggest some improvement in the mean test score following the instructions provided (and it is worth noting also that 19 of the 27 participants [70%] evidenced improvements in their post-attendance marks), the *t*-test procedure undertaken indicated no significant difference between these scores at the .05 level.

Pass rate was calculated as the number of tests passed divided by the total number of tests taken. For example, if a student passed three out of four tests that he or she sat, then his or her pass rate was 75%. In the university where this study was undertaken, pass rate is an often-used measure of whether a student is making satisfactory progress in his or her courses. A score of 50% or higher counted as a pass in the tests the students took.

The *t*-test undertaken showed that the participants' mean post attendance pass rate of 64.81% ($SD = 41.97$) was significantly higher than their mean pre-attendance pass rate of 51.41% ($SD = 39.31$), $t(26) = 2.50$, $p < .05$ (two-tailed). Hence the participants demonstrated an overall improvement in their pass rate subsequent to attending the extra tutorials provided.

The 169 other students who did not attend the extra tutorials served as the 'control' participants in that they received the regular instruction provided in the course, but not the mnemonic instruction provided in the extra tutorials. Hence, their test performance represented what would normally have been expected of students in the course—without the intervention provided in this study. It needs to be pointed out that an analysis of the data using a 2-factor ANOVA was not appropriate because the control participants did not receive any 'treatment' as such. Hence, no natural 'pre-instruction' and 'post-instruction' dichotomy could be applied to the control participant data.

The control participants showed no significant differences at the .05 level in their mean scores and mean pass rates *across the six tests administered*. (Their mean score across the six tests was 56.18%, and their overall pass rate was 74.35%.) This finding suggests a general stability in the students' test performance—in the absence of any intervention.

There were also no significant differences found at the .05 level between the motivation scores (intrinsic, extrinsic, and amotivation) of the participants and those of the 48 other students who did not attend the extra tutorials provided. The participants' mean scores were 56.15, 60.65, and 7.35 for intrinsic motivation, extrinsic motivation, and amotivation respectively, with the corresponding mean scores for the control group

being 51.19, 59.92, and 8.48. This suggests that the students who attended the extra tutorials were no more (or less, for that matter) academically motivated than the other students who did not attend.

Student evaluation of the tutorials

Twenty of the 27 participants completed and returned the questionnaire evaluating the usefulness of the extra tutorial sessions provided. The mean ratings they provided for the different sessions are shown in Table 2.

Table 2. Students' Mean Usefulness Ratings of the Instruction Sessions Provided

Instruction sessions	Mean ^a	SD
Hiragana and Katakana	4.16	.86
Vocabulary	3.90	1.10
Kanji	4.33	.82

^a On a scale of 1 to 5, where 1 = "not useful" and 5 = "very useful."

Eighteen of the 20 students who completed and returned the questionnaire attended the hiragana and katakana sessions. When asked to indicate which hiragana characters they found mnemonics helpful in remembering, 44% reported all the characters, while 33% circled anywhere from 12 to 18 of the characters (mean = 14.83 characters, *SD* = 2.79). Where the katakana characters were concerned, 33% of these participants reported having found mnemonics helpful in remembering all of them, while 50% circled anywhere from one to 28 of the characters (mean = 14.78 characters, *SD* = 8.32).

The participants were asked if they thought the extra tutorials helped their performance in the course assessments (the tests they had to sit for the course). Sixty-five percent (65%) circled "yes" on the questionnaire, while 20% circled "no," and 15% did not respond to the question.

The students were also given an opportunity on the questionnaire to provide any comments they would like to make that could be helpful to the researchers. Six of the participants wrote comments. One simply wrote "Good," while another commented on the hard work devoted by

the tutor (the second author) to the project. Two of the comments were requests: one for more tutorials offered at various times so that more people could attend, and the other for the speed of teaching to be slowed down (in the course itself) as there were students finding it difficult to keep up. The final two were compliments. One student wrote:

I found learning the Japanese alphabet using pictures and words was helpful for me to remember them. I am very grateful for the patience that [second author's surname]-sensei displayed. She needed it! And I liked her pictures and explanations. I now know how to read hiragana and katakana relatively better—a few hiccups—but kanji still gets me. どうもありがとうございます。

Another wrote:

I have performed poorly in this [course]—but this has no reflection on these extra tutorials. They were the [course's] only redeeming feature ... Thanks for your help. I enjoyed the mnemonics tutorials.

Discussion

The extra tutorials on mnemonic use provided in the present study were all highly rated in terms of their helpfulness by the students who attended. Furthermore, the majority of the students indicated that they believed the sessions helped their performance in the tests they sat.

Is there any basis though for their belief that these mnemonic tutorials were helpful? Although the improvement shown by the participants in their actual test scores was not statistically significant, the majority of the participants did evidence improvements in their post-attendance test scores. More importantly, however, when pass rates were examined, a significant improvement in the participants' post-attendance pass rate was found. This indicates an improvement in their capability to pass the tests following attendance of the tutorials on mnemonic use. Hence there are good reasons for the positive appraisals made by the participants.

There is a criticism frequently levelled at studies showing improvements in students' academic performance following instructional intervention: that those who choose to participate in such intervention programmes are probably better motivated and apt to show improvements in performance anyway—even without the interventions. Authors such as Mealey (1990) argue that there are strong links between motivation and

academic performance. In the present study, however, motivation as a possible contributing factor was also examined and, where academic intrinsic, extrinsic, and amotivation were concerned, no significant differences were found between the students who attended the extra tutorials and those who did not. There was therefore no evidence to suggest that the pass rate improvements shown by those who attended the tutorials were simply due to greater motivation.

A related question that could be raised is: If the participants were no more or less academically motivated compared to their control counterparts, why were their mean scores and pass rates lower? The extra tutorials provided were advertised as being aimed at those who may be experiencing some difficulties in learning the Japanese script. Thus, the students who volunteered to attend (the participants) would have been those performing relatively poorly in the course. The possible reasons for their poor performance in the first place are many (including perhaps inappropriate study strategies) and outside the parameters of the present study to investigate.

It could be argued that because the instructions provided in the present study covered the use of pictorial mnemonics in remembering not only kana script characters but also vocabulary words and kanji characters, it would be difficult to claim positive effects of instructions provided *specifically* on retention of kana script characters. The participants, however, were asked to evaluate the usefulness of each of the instructional sessions provided and all—including those dealing with the kana script characters—were rated highly by the participants as being useful in helping them remember. As noted earlier, more than three-quarters of the participants explicitly indicated on the questionnaire they completed that the mnemonics instructions helped in remembering either all or specific hiragana and katakana characters they identified.

It was necessary in the present study to include instructions on the use of mnemonics in remembering not just kana script characters but also vocabulary and kanji because, as noted earlier, the course tests that the students had to sit did not just assess mastery of kana but also covered vocabulary, grammar, and kanji. This is one significant limitation of undertaking research in a natural setting. Control over what is covered in regular class sessions and student assessments, as well as selection of experimental and control participants and scheduling, is very much limited. The important advantage, however, is that stronger claims about applicability of methods employed in real life settings (rather than just artificial laboratory settings) can be made. In the present study, it has

been shown that the use of pictorial mnemonics can be helpful to students who are really learning Japanese and have to recognize, recall, and reproduce the Japanese scripts.

For language teachers who still view the use of mnemonic methods with reservation and/or suspicion (not least because they are deemed “unnatural”), the present authors can only offer reassurance based on their observations that the majority of students take to the use of these techniques very well. None of the participants reported any difficulties in generating their own mnemonics during the vocabulary and kanji sessions, confirming previous research findings that there is nothing unnatural about the generation and use of mnemonic strategies for remembering various forms of information (e.g., Brooks et al., 1993; Hill et al., 1993; Manalo, 1999). In fact, some of the mnemonics the participants devised were very clever. For example, to remember らいしゅう (pronounced as /raɪʃuː/, and meaning “next week”), one participant came up with the mnemonic sentence “I do not want to *rush you* so *next week* would be fine.”

Some language teachers may also feel uncomfortable about the division of Japanese words at arbitrary points (e.g., *a-sago-han*, instead of the more appropriate *asa-gohan*) that could result from generating English- and other-language-derived mnemonic keywords. However, the present authors are by no means advocating the use of mnemonics for all new vocabulary words, rather for teachers to familiarise students with mnemonic strategies so as to enable them to use these strategies *when necessary*. Presumably, with an increasing vocabulary and a growing appreciation of word structures in Japanese, students would gradually become less reliant on mnemonics and other similar “coping” strategies as they progress through their studies. However, it is important that students are equipped with such strategies at the beginning of their language acquisition—when explaining for example the morphological components of words would more likely overwhelm and discourage most students—to instil a sense of manageability and control over the demands of this process.

For future research, one particular area that appears to warrant further investigation concerns the relative effectiveness of each of the pictorial mnemonics associated with each of the hiragana and katakana script characters. It appears that some are more helpful in remembering the script characters than others. A number of the participants verbally commented that pictorial mnemonics that have a *real* similarity to the script character in question (e.g., the “two eels” for remembering ん) are far

more helpful than those that do not have an inherent similarity and require a stretch of the imagination—plus considerable manipulation of the script character representation—to appreciate (e.g., “two friends watching the setting sun” for remembering せ, which is pronounced /se/). This is an aspect that educators who develop and construct these kinds of teaching materials may similarly wish to consider and possibly address. The importance of the similarity of the mnemonic association with the target information has also been noted by Gruneberg (1987/1997) where keywords as linkwords are concerned, and by Bellezza (1987, 1996) who explained this importance in terms of the bi-directionality of the association between the mnemonic cue and the target information. According to Bellezza, the association must operate not only in the direction of target information to mnemonic cue, but also in the opposite direction of mnemonic cue to target information. If the similarity between these is weak or artificial, then the association either or both ways is compromised.

Another area that ought to be investigated in future research is the usefulness of mnemonic strategies in learning abstract kanji script characters. In both the present study and the earlier-mentioned Lu et al. (1999) study, the kanji used were those visually similar to the meanings they represent or are concrete in nature (making it possible, with enough imagination, to make pictorial associations with their meanings). The greatest difficulty that most learners of kanji encounter, however, is in learning the more complex and often abstract symbols that do not lend any obvious connection to objects or shapes that can be imagined. Thus, the development of effective and efficient strategies would be of great assistance to learners here.

In conclusion, the findings of the present study provide some support to the main hypothesis posed: students taking a first year Japanese course who attend extra tutorials showing them how pictorial mnemonics could be used to remember Japanese script characters do evidence an improvement in their course assessment. This however needs to be qualified: a significant improvement was found in pass rates, but perhaps because of the low number of participants, the difference in pre- and post-instruction mean scores did not reach statistical significance despite the improvement shown. There are certainly clear indications in the present study pointing to the usefulness of employing pictorial mnemonic strategies in learning the script of another language in real classroom settings—not just where scores are concerned, but also from students’ appraisals of those strategies. This is but one study however,

with the limitations already noted. The present authors therefore hope to stimulate the interest of other researchers to conduct further investigations in this area.

Acknowledgements

The authors would like to thank Yukako Sunaoshi, Alamelu Badrinarayanan, Taeko Oya, Rosalie Smith, Yu Watanabe-Manalo, Marcus Henning, and Glenis Wong-Toi for their help.

Author Biographies

Emmanuel Manalo, PhD, is Director of the Student Learning Centre at the University of Auckland. His research interests include memory, learning disabilities, communication skills, and postgraduate education. He is the author of three books and over 30 research articles on these topics. He is President of the Association of Tertiary Learning Advisors in Aotearoa/New Zealand (ATLAANZ).

Satomi Mizutani, MA, is a Research Assistant in the Department of Applied Language Studies and Linguistics at the University of Auckland. Her main research interest is language teaching innovation. Her Master's thesis investigated the effects of goal setting on the language learning motivation of university students.

Julie Trafford, MSc, is a Senior Tutor at the Student Learning Centre at the University of Auckland. She is Co-ordinator of the Business Communication Skills Development Program and the Centre's Summer School Program. She is the author of three books as well as numerous papers on tertiary student instruction and support that have been presented at international conferences.

References

- Atkinson, R. C., & Raugh, M. R. (1975). An application of the mnemonic keyword method to the acquisition of a Russian vocabulary. *Journal of Experimental Psychology: Human Learning and Memory*, 104, 126-133.
- Baddeley, A. D. (1997). *Human memory: Theory & practice* (Rev. ed.). Hove, East Sussex: Psychology Press.
- Bellezza, F. S. (1987). Mnemonic devices and memory schemas. In M. A. McDaniel & M. Pressley (Eds.), *Imagery and related mnemonic processes* (pp. 34-55). New York: Springer-Verlag.

- Bellezza, F. S. (1996). Mnemonic methods to enhance storage and retrieval. In E. L. Bjork & R. A. Bjork (Eds.), *Memory* (pp. 345-380). San Diego: Academic Press.
- Brooks, J. O. III, Friedman, L., Gibson, J. M., & Yesavage, J. A. (1993). Spontaneous mnemonic strategies used by older and younger adults to remember proper names. *Memory*, 1 (4), 393-407.
- Brown, J. D. (1992). Statistics as a foreign language—Part 2: More things to consider in reading statistical language studies. *TESOL Quarterly*, 26, 629-662.
- Condu, M. M., Marshall, K. J., & Miller, S. R. (1986). Effects of the keyword mnemonic strategy on vocabulary acquisition and maintenance by learning disabled children. *Journal of Learning Disabilities*, 19, 609-613.
- Desrochers, A., Gelin, C., & Wieland, L. D. (1989). An application of the mnemonic keyword method to the acquisition of German nouns and their grammatical gender. *Journal of Educational Psychology*, 81, 25-32.
- Ehri, L. C., Deffner, N. D., & Wilce, L. S. (1984). Pictorial mnemonics for phonics. *Journal of Educational Psychology*, 76, 880-893.
- Elliott, J. L., & Gentile, J. R. (1986). The efficacy of a mnemonic technique for learning disabled and nondisabled adolescents. *Journal of Learning Disabilities*, 19, 237-241.
- Ellis, N. C., & Beaton, A. (1993). Factors affecting the learning of foreign language vocabulary: Imagery keyword mediators and phonological short-term memory. *Quarterly Journal of Experimental Psychology*, 46A, 533-558.
- Gruneberg, M. M. (1987/1997). *Linkword French, German, Spanish, Italian, Greek*. London: Corgi Books/Lincolnwood N.T.C./Nailsea Interaktive Software.
- Gruneberg, M., & Sykes, R. (1996). The use of mnemonic strategies in the learning of non-roman foreign language alphabets. *Language Learning Journal*, 13, 82-83.
- Heerman, C. E., & Maleki, R. B. (1994). Helping probationary university students succeed. *Journal of Reading*, 37, 654-661.
- Hill, R. D., Schwob, S. L., & Ottman, S. (1993). Self-generated mnemonics for number recall in young and old adults. *Perceptual and Motor Skills*, 76, 467-470.
- Levin, J. R. (1983). Pictorial strategies for school learning: Practical illustrations. In M. Pressley & J. R. Levin (Eds.), *Cognitive strategy research: Educational applications* (pp. 213-237). New York: Springer-Verlag.
- Levin, J. R., McCormick, C. B., Miller, G. E., Berry, J. K., & Pressley, M. (1982). Mnemonic versus nonmnemonic vocabulary-learning strategies for children. *American Educational Research Journal*, 19, 121-136.
- Levin, J. R., & Pressley, M. (1985). Mnemonic vocabulary instruction: What's fact, what's fiction. In R. F. Dillon (Ed.), *Individual differences in cognition* (Vol. 2). Orlando, FL: Academic Press.

- Lu, M., Webb, J. M., Krus, D. J., & Fox, L. S. (1999). Using order analytic instructional hierarchies of mnemonics to facilitate learning Chinese and Japanese kanji characters. *Journal of Experimental Education*, 67 (4), 293-311.
- Manalo, E. (1991). The incorporation of process mnemonic instruction in teaching computational skills: A case report on a mathematics learning disabled individual. *Focus on Learning Problems in Mathematics*, 13 (4), 21-34.
- Manalo, E. (1997). Applications of mnemonics in education: A brief review. In G. M. Habermann (Ed.), *Looking back and moving forward: 50 years of New Zealand psychology* (pp. 157-166). Wellington and Palmerston North: The New Zealand Psychological Society.
- Manalo, E. (1999). Spontaneous mnemonic use in simulated foreign word learning. *Psychologia - An International Journal of Psychology in the Orient*, 42, 160-169.
- Manalo, E., Bunnell, J. K., & Stillman, J. A. (2000). The use of process mnemonics in teaching students with mathematics learning disabilities. *Learning Disability Quarterly*, 23 (2), 137-156.
- Manalo, E., Wong-Toi, G., & Henning, M. (1996). Effectiveness of an intensive learning skills course for university students on restricted enrolment. *Higher Education Research and Development*, 15 (2), 189-199.
- Mastropieri, M. A., Scruggs, T. E., & Fulk, B. J. M. (1990). Teaching abstract vocabulary with the keyword method: Effects on recall and comprehension. *Journal of Learning Disabilities*, 23, 92-96, 107.
- Mastropieri, M. A., Scruggs, T. E., Levin, J. R., Gaffney, J., & McLoone, B. (1985). Mnemonic vocabulary instruction for learning disabled students. *Learning Disability Quarterly*, 8, 57-63.
- McDaniel, M. A., & Pressley, M. (1989). Keyword and context instruction of new vocabulary meanings: Effects on text comprehension and memory. *Journal of Educational Psychology*, 81, 204-213.
- McDaniel, M. A., Pressley, M., & Dunay, P. K. (1987). Long-term retention of vocabulary after keyword and context learning. *Journal of Educational Psychology*, 79, 87-89.
- Mealey, D. L. (1990). Understanding the motivation problems of at-risk college students. *Journal of Reading*, 33, 598-601.
- Ogawa, K. (1990). *Kana can be easy*. Tokyo: The Japan Times Ltd.
- Plater, W. M., Matthews, J. R., D'Appollonia, S., & Abrami, P. C. (1997). In response ... *Change*, 29 (5), 16-18.
- Quackenbush, H. (1999). *Hiragana/katakana in 48 minutes: Teacher guide*. Carlton, South Victoria: Curriculum Corporation.
- Quackenbush, H., Chujo, K., Nagamoto, K., & Tawata, S. (1989). Teaching how to read hiragana in 50 minutes: A comparison of mnemonics and the use of cards with associated colours. *Journal of Japanese Language Teaching*, 69, 147-162.

- Quackenbush, H., & Ikeda, S. (1989). *Katakana in 48 minutes*. Canberra: Curriculum Development Centre.
- Quackenbush, H., & Ohso, M. (1983). *Hiragana in 48 minutes*. Canberra: Curriculum Development Centre.
- Richardson, J. T. E., Cermak, L. S., Blackford, S. P., & O'Connor, M. (1987). The efficacy of imagery mnemonics following brain damage. In M. A. McDaniel & M. Pressley (Eds.), *Imagery and related mnemonic processes* (pp. 303–328). New York: Springer-Verlag.
- Rosenheck, M. B., Levin, M. E., & Levin, J. R. (1989). Learning botany concepts mnemonically: Seeing the forest and the trees. *Journal of Educational Psychology*, 81, 196–203.
- Takabe, Y. (1993). *Kanji isn't that hard! Kanji can be mastered with the "24 Rules."* Tokyo: ALC Press.
- Trout, P. A. (1997). What the numbers mean. *Change*, 29 (5), 24–30.
- Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C., & Vallieres, E. F. (1992). The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement*, 52, 1003–1017.
- Walsh, R. W. (1985). Changes in college freshmen after participation in a student development program. *Journal of College Student Personnel*, 26, 310–314.
- Young, S., & Shaw, D. G. (1999). Profiles of effective college and university teachers. *Journal of Higher Education*, 70, 670–686.

Perspectives

English in Japan: The World Englishes Perspective

Phillip R. Morrow

Nagoya Gakuin University

This paper examines the role of English in Japan from the World Englishes (WE) perspective, concentrating on two issues: the implications of WE for English education, and the status of Japanese English (JE) as a variety of English. An overview of WE is followed by a discussion of its implications for English teaching in Japan. Important implications include the need to familiarize students with multiple varieties of English and to encourage them to regard all varieties, including their own, as valid. In this connection, the status of JE is discussed and research findings are cited to support recognizing JE as an independent variety of English.

本論文では、“World Englishes”(WE)の観点から、二つの関連した問題を検証する。すなわち、WEが日本の英語教育に対して果たす役割、および、英語の一形態としての“Japanese English”(JE)の位置づけである。まず、WEの理論を概観した上で、WEが日本の英語教育において果たす役割を論じる。次に教育において英語のどの形態を用いるべきかということについて、学生に多様な形態に親しませること、および、自分たちの英語を含めて、様々な英語の形態が正当なものであることを学生に気づかせることが重要であることを論じる。これに関連してJEの位置づけを行い、JEは多様な形態をもつ英語の一つとして認めるという主張を裏付ける研究結果について論じる。

The use of English in places where it is not a native language has been the object of much recent research by linguists (e.g. Crystal, 1997; Jenkins, 2000; Thumboo, 2001), and the area of linguistics most directly concerned with the description of these varieties and their users and uses has come to be known as *World Englishes* (hereinafter, WE). The aim of this paper is to examine the role of English in Japan from the WE perspective. It will deal mainly with two issues. The first is the implications of the WE approach for the teaching of English in Japan. The second concerns the status of Japanese English (hereinafter, JE) as a variety of English. I will argue that the English used by Japanese native speakers has characteristics and usages that distinguish it from other varieties of English and justify regarding it as an independent variety of English. As evidence for this position, I will cite the findings of a number of discourse-level studies of the English used by Japanese speakers. By way of introduction to these issues, I will briefly review some pertinent facts about the spread and use of English and the WE perspective.

The Spread and Use of English

English occupies a unique place in the world today and in history. There has never been a language which has been used so much by so many different people. David Crystal, in his book, *English as a Global Language* (1997), lists 75 countries where English is used. In some 47 of them, the non-native speakers of English outnumber the native speakers, and in many of those countries there are almost no native speakers of English. In these places, English often has the role of a *lingua franca*, that is, a language of communication among people who do not speak the same native language. English is not the first language to serve as a *lingua franca*, of course. Since antiquity there have been languages that have had this function at different times and in different parts of the world, for example, Greek, Latin, Sanskrit, French, and Hausa. In each case, the choice and use of a language as a *lingua franca* has been related to political factors, and English is no exception. The position of English in the world today can be seen as one legacy of British colonial policy.

What is exceptional about English is the extent to which English has come to be used. There are now five languages in the world that have a very large number of speakers: Chinese, English, Hindi-Urdu, Russian and Spanish. But among these, only English can claim to be a real universal language, that is, a language used for communication between

peoples of diverse cultural and linguistic backgrounds. In this sense, English is unique.

Today the position of English as an international language is unrivaled. It is now the official or semi-official language in more than 60 countries. Moreover,

English has become the dominant language in many fields of activity such as business and banking, industry and commerce, transportation, tourism, sports, international diplomacy, advertising, pop music and so on. But above all, English has become the common language of scientific discourse in a world where the relative “development” of a nation can best be measured in terms of its access to science through English. (Medgyes, 1994, p. 1)

It is very difficult to estimate the number of English users with accuracy; after all, who counts as a user of English? Is it a matter of education? Or is it a matter of regular use? If having studied English for some years in school qualifies a person as an English user, then there would be about a hundred million users of English in Japan, but many of these people cannot be said to “use English” in any meaningful way. Thailand is similar to Japan in being an Asian country that was not colonized by any European power. In Thailand, as in Japan, English is considered a foreign language rather than a second language, but unlike Japan, a large number of ordinary people—shop attendants, clerks, office workers, and so on—are able to use English and do use English for at least some functions in the course of their daily lives. Many of these people have had little formal instruction in English, yet they use English regularly, often because their work brings them into contact with tourists, and tourism is a major industry in Thailand. Should these people be considered English users? There is no widely accepted criterion for deciding who is a user of English, and this accounts for the great variation in estimates about the number of users of English.

Crystal (1997) puts the number of English L1 speakers at about 337 million and the number of English L2 speakers at 235 million as of 1995. These speakers come from countries whose populations total more than 2 billion, or more than one third of the world’s population. In theory, most of these 2 billion people are routinely exposed to English, and besides the English L1 and L2 speakers that are included in these figures, some others have varying degrees of proficiency in English. For instance, in India only about 4% of the population are counted as English L2 speakers in Crystal’s figures, but if anyone who uses English for any function

were included, the figure would double or treble. Furthermore, these figures do not include the EFL speakers in countries like Japan, Taiwan, or Thailand. Whatever criterion is used for counting users of English, the fact remains that the number of English speakers is very large, and it is increasing with the spread and use of technology.

The World Englishes Perspective

World Englishes is the term that has come to be used to designate the area of research that encompasses the issues related to the spread and use of English throughout the world. These issues fall within many of the traditional areas of linguistics: Sociolinguistics, Dialectology, Language Contact, Bilingualism, Language Education, Language Planning, and others. Clearly, these are very diverse areas and the WE studies that deal with issues in these various areas are not all based on a shared theory. Rather, they are based on a few common premises that have important linguistic, pedagogical, and social implications. One of these premises is that English has become a global language. It is learned and taught and used to an extent unprecedented in history. Another basic premise is that there have been significant demographic changes in terms of who the users of English are, and where they are. A third important premise is that there are many, many varieties of English, and these varieties can be distinguished from one another in terms of their phonological, lexical, syntactic, and discursive characteristics.

Kachru (1985) has proposed a useful scheme for classifying the areas of the world where English is used. He divides the English-using world into three concentric circles: an Inner Circle, an Outer Circle, and an Expanding Circle. (See Figure 1 below.) The Inner Circle consists of the countries where English is learned and used as a first language by most of the inhabitants. Australia, Canada, New Zealand, the UK, and the USA are examples of countries in the Inner Circle. The Outer Circle comprises mostly the former colonies or spheres of influence of the UK and the USA. This includes: Bangladesh, Botswana, Cameroon, Ethiopia, Fiji, Gambia, Ghana, India, Israel, Kenya, Lesotho, Liberia, Malawi, Malaysia, Malta, Nigeria, Pakistan, the Philippines, Sierra Leone, Singapore, South Africa, Sri Lanka, Sudan, Tanzania, Tonga, Uganda, Western Samoa, Zambia, and Zimbabwe, among others. In these countries, nativized, that is, local varieties of English have achieved the status of being either an official language, or a language widely used in education, administration, law, business, the mass media, and literature. The Expanding

Circle consists of countries where English is on the way to becoming a dominant second language in the domains of education, science, and technology. These countries include, among others, China, Japan, Taiwan, Thailand and the countries of Europe.

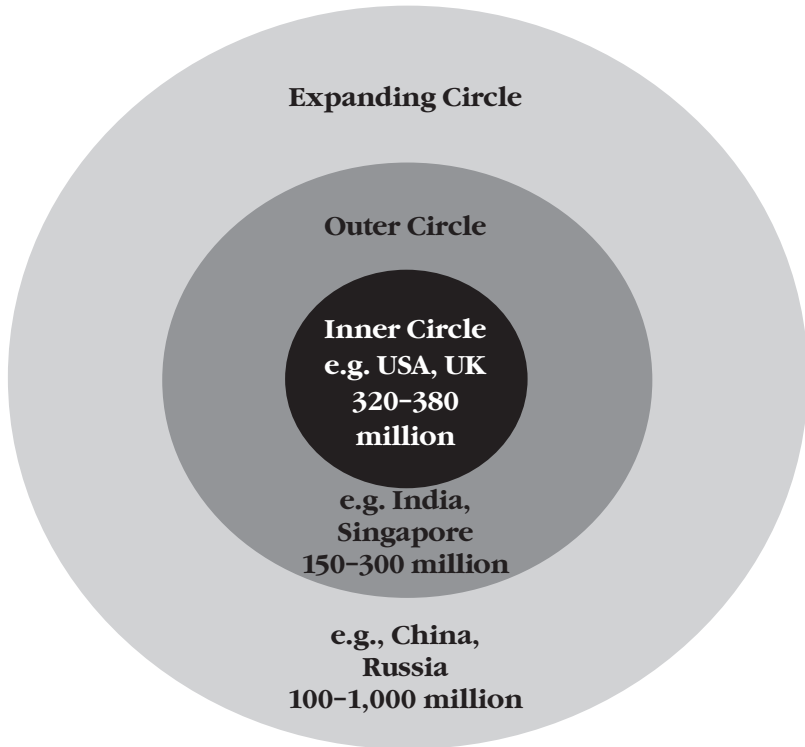


Figure 1. The three circles of English (from Crystal, 1997, p. 54)

With the spread of English there has been a significant change in the demographic distribution of its users. The proportion of English users from the Outer and Expanding Circles has become larger and is increasing. There are, for example, more English speakers in India than in Australia and New Zealand combined. An important implication of this is that more and more users of English today are bilinguals or multilinguals for whom English is a second, or a third language.

The spread of English has also fostered the emergence and development of new varieties of English. There are now many varieties of English, each of which reflects the cultural conditions of the place or places where it is used. The existence of these varieties raises a question about their relative ranking. Are some varieties better or purer than others? Linguists do not believe that any language or variety is inherently better than any other, though they recognize that varieties and languages may differ as to the size of their lexicons and the functions for which they are used. However, languages and varieties are not all equal in the eyes of their users. There are large differences in the levels of prestige that accrue to different varieties, and these differences are related to many factors including the functions for which the variety is used and the region with which it is associated. For example, the varieties of the national language used in the political capitals such as Paris, London or Tokyo tend to have more prestige than the local varieties of the same languages that are used in different areas of the countries, although there is no linguistic reason for this. Rather, this reflects the tendency that languages or varieties used by those with greater status and power are more highly valued than those used by speakers with less status and power (Romaine, 2000).

In the Outer Circle and the Expanding Circle, the extent to which English is used and the functions for which it is used vary greatly from one place to another. In WE studies the concepts of *range* and *depth* are used to describe the different dimensions of the spread and use of English. According to Kachru, “The *range* of a variety refers to its extension into various cultural, social, educational and commercial contexts....The term *depth* relates to the penetration of bilingualism into various strata of society” (1986, p. 92). Saying that a variety has wide range means that it is used in a number of areas such as commerce, religion, and law. If a variety has depth, then it is used by individuals from many different levels of society, for example, government officials, scholars, professionals, merchants, and technicians.

Among the varieties of English in the Outer and Expanding Circles, a general distinction is made between Institutionalized Varieties and Performance Varieties. Institutionalized Varieties (e.g. Indian English, South African English) are those used in places where English has some sort of official status, either in law or through established and recognized usage. Institutionalized Varieties are generally associated with the Outer Circle, and for those who use these varieties, English is a second language. Performance Varieties (e.g. Korean English, Thai English) are

varieties used in places where English has not received any kind of official recognition. They are usually associated with the Expanding Circle, and for their users, English remains a foreign language.

The distinction between Institutionalized and Performance Varieties becomes clearer if we look at individual cases. Indian English is a good example of an Institutionalized Variety (D'souza, 2001, Kachru, 1986). English in India has a long history dating back to 1600 when the East India Company was formed for purpose of trade with India. Since then, and particularly during India's colonial period, the use of English became firmly entrenched. After independence, English was given the legal status of "associate official language" and has come to be widely used in the media, in education, in commerce, in technology and in other domains. It is also used as a *lingua franca* among Indians who have different native languages. The variety of Indian English used by many speakers has certain phonological characteristics and lexical and grammatical usages (described in Kachru, 1983) that set it apart from other varieties. English is used in some states of India much more than in others, and there is considerable variability in the level of English proficiency among individuals, but it is fair to say that English has considerable range and depth in India.

The situation in Korea is quite different: "...in Korea English has never been institutionalized, nor has it been a(n intra)national or official language" (Song, 1998, p. 263). Song adds that although most Koreans have studied English at school, the majority are unable to carry out simple communicative functions. While Korean English clearly has much less range and depth than Indian English, Korean English does have a number of lexical, morpho-syntactic and pragmatic usages, which distinguish it from other varieties of English (Shim, 1999). These usages have been codified by being included in the English textbooks used in Korea. In WE studies, the distinctive usages that set Korean English (and other performance varieties) apart from Inner Circle varieties of English are termed *deviations* rather than *mistakes* since they reflect the norms of the linguistic and cultural setting in which they are used, and are therefore acceptable usages within certain contexts.

Studies of the many varieties of English have drawn attention to the diversity of the language. This is an aspect of English which has not been emphasized much in the teaching of English, and yet would seem to have important implications for learners. In the following section some of these implications are considered in relation to the case of TEFL in Japan.

World Englishes and the Teaching of English

Research on WE has raised awareness of how English is used in non-native contexts, and has shown that non-Inner Circle varieties are functionally adequate and valid as varieties of English. For EFL teachers, the existence of multiple varieties of English raises a question: Which variety should be used as a model, and how should other varieties be treated in the classroom? For many teachers, there is an obvious answer: Standard English.

Standard English

The adoption of Standard English (hereinafter, SE) as the model for EFL teaching is not quite as straightforward and unproblematic as it might seem. One large problem has to do with the very notion of SE, as Widdowson (1996) has shown in a paper entitled, "The Ownership of English." SE has usually been defined in terms of its grammar and lexis, but Widdowson believes that a standard lexis does not exist. He points out that there are many technical words that are used in newspaper articles and yet are not found in dictionaries. Surely those words must be considered part of SE. Similarly, the grammatical usages which are supposed to set SE apart from other varieties of English are rather slight, and not so great as to impede communication. Is SE really different then from other varieties in any significant way? Widdowson thinks not. Rather, SE is a variety, a kind of superposed dialect, which is socially sanctioned for institutional use and, therefore, particularly well suited to written communication (Widdowson, 1996).

Widdowson notes that the idea of a standard language implies stability and this is an attractive feature for a language variety that is going to be used as a model for pedagogical purposes. But a language is a living thing, and like all living things it is continuously growing and changing. It is, therefore, inherently unstable. There are always new words coming into use and others falling into disuse. The innovating aspect of language is a natural one and it comes into play when a language is used in a new cultural context or when users have a need to develop ways to talk about new technologies or other new things. There is always some innovation going on at the edges of the language. This is the normal state of affairs, and linguistic innovation is something that speakers admire. Using words in new and unexpected ways is one expression of creativity in the poems of the great poets and in the speeches of skilled orators. But, as Widdowson (1996) points out, creativity with words is not admired

nor even permitted in the case of EFL learners. Learners are expected to learn the rules and usages of SE and adhere to them strictly.

Widdowson is not the only one to express doubts about SE. Nayar (1997) and Davies (1999) have also questioned the validity of SE, and even among those who, like Widdowson, accept the existence of an SE variety the definition of SE is controversial. According to *The Concise Oxford Companion to the English Language* (McArthur, 1996) "Standard English" is,

A widely used term that resists easy definition but is used as if most educated people nonetheless know precisely what it refers to. Some consider its meaning self-evident: it is both the usage and the ideal of 'good' or 'educated' users of English. A geographical limitation has, however, often been imposed on this definition, such as the usage of educated people in Britain alone, England alone, of southern England alone, or the usage of educated people in North America and Britain generally. Others still find standard English at work throughout the English-speaking world. (p. 902)

Thus, there are multiple and conflicting definitions of "Standard English." But if we discard SE as the model for language teaching, where does that leave us? What can replace it? D'souza (1999) has addressed this point:

The emphasis in the total English curriculum should be on diversity, richness, variation rather than on sameness and uniformity. ... Students must be given the maximum exposure to written and spoken texts so that they realize that they can expect to encounter Englishes that differ from their own and are equipped to deal with them. Advocates of a single monolithic standard never fail to raise the bogey of 'intelligibility' as one of the arguments against acceptance of multiple standards. What we must keep in mind is that intelligibility depends in large part on familiarity. Therefore exposing all speakers of English to as many varieties of English as possible would do more to insure intelligibility than trying to impose a single standard on everyone. (p. 273)

This view receives support when we look at the needs of Japanese users of English. Some years ago I conducted a survey of some 20 Japanese businessmen who were in a one-year training program for international business at an American university. One finding was that they most often used English not for interacting with native speakers, but for interacting with other non-native speakers. Their needs would therefore be better served by exposure to many varieties of English rather than just SE.

Native and non-native speakers

When the subject of different varieties of English is broached, the difference between native speakers and non-native speakers is still often stressed. This brings up another point: the dichotomy between native and non-native speakers. While it may be possible to identify native speakers and non-native speakers in speech communities that are basically monolingual, it becomes very hard to say who is a native speaker in multilingual communities. Davies (1995) has proposed five defining features of a native speaker: a) childhood acquisition, b) intuitions about one's idiolect and the standard language, c) ability to make fluent spontaneous discourse, d) potential for creativity, and e) unique ability to translate into L1. But Nayar (1997) counters that,

...none of these are really necessary or invariably shown to be present in all average speakers of any Inner Circle variety. Besides, anyone, with some fluency and the right citizenship, without fulfilling any of these conditions can claim to be a native speaker while someone who fulfils all the above may still be denied the status on ethnopolitical or domicile reasons. (p. 286)

Kachru (1985) also believes that the native/non-native speaker dichotomy has become irrelevant and other sociolinguists hold this view, as well. Ferguson (1992) has written,

Some languages ... spread widely as lingua francas between speakers of different languages or serve as languages of special functions in communities of non-native speakers; this kind of language use merits the attention of linguists as much as do the more traditional objects of research. In fact, the whole mystique of native speaker and mother tongue should probably be quietly dropped from the linguists' set of professional myths about language. (p. xiii)

Whether or not the concept of "native speaker" has validity is a question of some moment for EFL teachers, because being a native speaker has been counted as an important—and sometimes, required—qualification for employment. But if the validity of the concept of a native speaker is dubious, then the requirement that EFL teachers be native speakers must also be questioned. From the WE perspective, it is highly desirable to expose students to different varieties of English, and for this reason it is desirable to employ not only English teachers who are speakers of one of the Inner Circle varieties, but teachers who are speakers of one of the Outer Circle or Expanding Circle varieties, as well.

Teaching Varieties of English

In Japan, there are a few teachers from India, but very few, and there are almost none from any other Outer Circle or Expanding Circle country. This is unfortunate, as having a teacher who speaks an Outer Circle or an Expanding Circle variety could be a positive role model for Japanese students. It would help to make students aware that English is not something that belongs only to native speakers or is used only when speaking to someone from an Inner Circle country. Those responsible for English education should prepare students to interact with users of English whose variety of English may be very different from the students' own. Communicative success in such encounters may depend in part on having a positive attitude toward the language of others who speak a different, but equally legitimate variety of English. One way of inculcating this attitude is to promote contact with speakers and with teachers from the Outer Circle. This could have the additional merit of encouraging students to see their own variety as a valid one.

Recognizing the need to familiarize students with multiple varieties of English is a first step toward incorporating the WE perspective into EFL teaching, but there remains the problem of locating teaching materials that present multiple varieties of English, particularly varieties from the Outer Circle and the Expanding Circle. Most EFL textbooks, including those published in Japan, take Standard American English or Standard British English as the model. Some listening materials, for example, those in the *Cambridge Skills for Fluency* series do present alternative dialects of English, although in this series the alternative dialects of English all seem to be other Inner Circle dialects (Doff, 1993). Jenkins (2003) provides a comprehensive introduction to key topics and present issues in the area of WE. Each chapter includes discussion questions and activities that could be used with advanced EFL students to increase their awareness of the different varieties of English and the issues related to the use of these varieties. Aside from textbooks there are a few useful sources, which could be used to advantage in acquainting students with different varieties of English. One in particular is *The Story of English*, a book (McCrum, Cran & MacNeil, 1986) and video series (McCrum, Cran, MacNeil, & Pett, 1986), which portrays the history and spread of English. This series has the merit of showing very clearly the variation, which is so much a part of the language, in all of its varieties.

Another reason that students do not learn much about different varieties of English is that many teachers are not themselves too knowledge-

able about varieties of English, especially the non-Inner Circle varieties. Y. Kachru (2001) has written that,

...a perspective on language variation and its implications for language learning and teaching...is almost totally absent in SLA ('second language acquisition') literature. There is resistance to acknowledging the social reality of varieties and their relevance for human interaction across languages and cultures. (p. 349)

This is reflected in a lack of courses about language variation and World Englishes in teacher training programs for EFL teachers. The WE perspective has hitherto been a minority one. It is only minimally reflected in current TEFL methodology texts or in the most popular education journals. In 1993, Brown reported that in three popular methods texts (Celce-Murcia, 1991, Brown, 1987, and Long & Richards 1987), there was very little about WE, and that it was sometimes left out altogether. Ten years later, a cursory look at some representative texts used for EFL teacher training (e.g. Brown, 2000, Cook, 2001, Harmer, 2001 and Ur, 2001) indicates that the situation remains little changed. Thus, at this point, the inclusion of the WE perspective remains an individual effort.

Japanese English: A Variety of English?

When EFL instructors in Japan attempt to integrate the WE perspective into their teaching, one central issue that they must consider concerns the status of the English used by Japanese L1 speakers. Is there an identifiable variety of English that can be called Japanese English?

In Japan, English is clearly the number one foreign language. Almost all junior high school and high school students study English (Koike & Tanaka, 1995). But although it is the most studied foreign language, its use within Japan is quite limited in range and depth. Regarding range, Yano (2001) observes that there are very few domains or functions in which English has supplanted Japanese, nor even where English is an alternative. And as for depth, the numbers and types of users who have occasion to interact with others in English are small, though there are an increasing number of users in scientific or technical fields who read English materials for professional purposes and there are others who use English for business-related correspondence.

Those who have used the term *Japanese English* have usually done so in the context of discussions of English loanwords, and numerous research articles deal with the topic of loanwords in Japanese (e.g. Kay,

1995, Morrow, 1987, Stanlaw, 1982). There are large and increasing numbers of loanwords from English in Japanese, and the spread and use of these loanwords furnishes interesting data for a study of language contact. Loanwords are pronounced following the rules of Japanese phonology, they often take on new and quite different meanings, and they are used for certain discourse functions, such as making something sound fashionable (Takashi, 1990). Because of their distinctive phonological forms, meanings and usages they have been referred to as *Japanese English* in studies of loanwords, but the use of loanwords cannot be considered strong evidence for a Japanese variety of English since the loanwords are used in Japanese discourse, not English discourse.

Better evidence for JE as an independent variety of English is available from studies of Japanese speakers' spoken and written English. There are now a number of studies that have identified discourse level features that distinguish JE from other varieties of English. From a study of Japanese and American ways of participating in business meetings, Yamada (1990) found some noteworthy differences in turn-taking. The Japanese in her study took short turns and shared turns fairly equally—that is, each speaker took roughly equal numbers of turns regardless of who initiated a topic. In contrast, the Americans took long monologic turns, and distributed their turns unevenly, with the participant who initiated a topic taking the highest proportion of turns in that topic. This indicates a difference in turn-taking rules between JE and American English (hereinafter, AE).

Another discourse characteristic of JE is a high frequency of backchanneling, the verbal and non-verbal signals that a listener produces to indicate that he/she is attending to what the speaker is saying. White (1989) measured the frequency of five verbal backchannel expressions (*mmhm, yeah, oh, uh huh, hmm*) and found that Japanese listeners in English conversations used these expressions far more than American listeners did. Similar results were obtained by Maynard (1997), who studied backchannel behavior as one aspect of interactional management. Her analysis was based on videotaped conversations between four American native English speakers and four Japanese non-native speakers of English, who were all attending an American university. By using videotaped conversations, she was able to include both verbal and non-verbal backchanneling in the analysis. The Japanese speakers of English in her study used backchanneling much more frequently than the American speakers did. Furthermore, there were differences in the type and placement of the backchannel responses used by the

Japanese and American speakers. The Japanese speakers used head movement (nodding and head shaking) as the most frequent type of backchannel response, whereas for American speakers brief utterances were the primary device. American speakers did not make backchannel responses during phrasal units, but waited until the pause after the phrasal unit. Japanese speakers, on the other hand, gave backchannel responses within the phrasal units while their interlocutor was speaking. Maynard concluded that these differences in backchannel behavior reflected differences in “interactional management” between Japanese and American speakers.

There is evidence of other types of interactional differences, as well. Murata (1995) noted a difference between Japanese speakers of English and native English speakers in the use of repetition, especially hesitation repetition. She suggested that one possible explanation for this is that quick turn-taking may appear to be too aggressive and intrusive to Japanese speakers and that they seek to lessen the intrusiveness and show respect for the “territoriality” of their conversation partner by using repetitions. Murata also observed a special tendency for Japanese speakers to use repetition for reformulation, and hypothesized that it was because Japanese speakers were “error-conscious” when they were speaking English (Murata, 1995, p. 352). In studies of topic maintenance (e.g. Chaudron & Parker, 1990, Sasaki, 1997) Japanese speakers of English have been found to have distinctive strategies of topic maintenance, particularly in the kinds of nouns and pronouns that they use to maintain topics. There may also be differences in the frequency of use of different categories of lexical items. Suenobu, Yamane & Kanzaki (1997) found that Japanese speakers and native English speakers used different proportions of content and function words. Yano suggests that Japanese speakers have a preference for the passive voice, saying, for instance, “The plan was decided on,” rather than, “We decided on the plan” (2001, p. 127).

Pitch is another distinguishing feature of JE, at least among male speakers. Loveday (1981) made a study of Japanese speakers speaking in Japanese and in English. He compared the pitch levels that these speakers used with the pitch levels used by British English speakers. In this study, male Japanese speakers of English used different pitches from those of male British English speakers. Physiologically, Japanese males and British males have about the same pitch range, but the top pitch used by the Japanese males was much lower than the top pitch used by the British males. Japanese males used lower pitch for greetings, goodbyes, and thank you’s than British males. This was true when they

spoke in Japanese, and also when they spoke in English. For instance, in thanking someone for a dinner invitation, all the British males raised their pitch, but the Japanese males did not.

From a study of pragmatics, there is evidence that Japanese speakers of English carry out some speech acts differently from American English speakers. Beebe, Takahashi and Uliss-Weltz (1990) found that the refusals of Japanese speakers showed evidence of negative transfer in the order, frequency and content of elements. For instance, in refusing a request to a lower status person both American and Japanese speakers tended to give excuses, but the Japanese speakers gave the excuse second, after an expression of positive opinion or an expression of regret, while the American speakers tended to give both a statement of positive opinion and an expression of regret before the refusal itself. In content, the excuses given by the Japanese speakers were more vague. For example, one Japanese speaker refused a request saying, "My children have many problems." Another study of refusals by Maeda (1989) found variation between the refusals of Japanese speakers of English and those of American speakers. Interestingly, she found that Japanese speakers used less politeness than American speakers, but she noted that variation depended on demographic factors, especially age and sex. Teenagers used the least polite forms and males used less polite forms than females. In particular, she found that Japanese males used far fewer politeness strategies than American speakers when they refused a request made by a lower status person or family member.

The findings of the studies reviewed here indicate that at the discourse level, Japanese speakers of English are operating with a set of rules that differs in many respects from those of speakers of Inner Circle varieties of English. Thus, these findings provide strong support for accepting JE as an independent variety of English. This does not mean that JE is or will become an institutionalized variety like Singaporean English or Indian English; rather, it means that we are justified in considering JE a performance variety.

More such findings that show the distinctiveness of JE may be forthcoming soon from another source. A group of researchers has been compiling a Japanese EFL Learner Corpus (JEFLC) which, when completed, is to include samples not only from advanced learners, but also from novice and intermediate learners. At present, they have collected some 200,000 words of both written and spoken (but mostly written) data.¹ Such corpora of actual usage of English by Japanese speakers will furnish a basis for the description of JE.

This brings us to the question of acceptance of JE and other non-Inner Circle varieties. In Japan, this issue is linked to a larger, political one: the acceptance and use of English itself. There are, as Kubota (1998) has pointed out, conflicting ideologies of English in Japan. On the one hand, there is a trend toward *kokusaika* 'internationalization' in government and business, and this ideology has stressed learning English as a means of making the Japanese nation and people more "international." On the other hand, another prevailing ideology is *nihonjinron*, which "attempts to define a distinct Japanese cultural and linguistic identity vis-à-vis the Western culture and language: particularly English" (Kubota, 1998, p. 299). This ideology, which emphasizes the cultural uniqueness of Japan, has often been cited by scholars and business leaders who see it as a major cause of Japan's economic success in the 1960s and 1970s (cf. Sugimoto & Mouer, 1980 [cited in Kubota, 1998]). This view does not encourage the teaching and use of English, and some (e.g. Tanaka, 1993 [cited in Kubota, 1998]) even see the teaching and use of English as a form of colonization. They fear that the spread and use of a foreign language (in this case, English) could diminish the role of the national language, and in this way threaten Japan's distinctive culture. This is an extreme view, and there are of course many more nuanced positions between those opposed to the teaching and use of English and those who welcome it wholeheartedly. The fact that English is included in the curricula of almost all secondary schools can be seen as an indication of a general and official acceptance of English, but some groups and individuals have ambivalent views toward English, and others who are in favor of the teaching and use of English may be unwilling to accept varieties of English from outside the Inner Circle. Toward these varieties, there is resistance from the educational establishment and from many learners themselves.

Some influential scholars (e.g. Quirk, 1995) feel that it is essential to have a single standard variety to serve as the norm for pedagogical purposes in EFL contexts, and that if multiple varieties of English are accepted, it will lead to confusion and problems of intelligibility between speakers of different varieties. For EFL teachers, expertise in Standard English has been counted as a valuable professional qualification, and teachers may feel that if alternative varieties of English are introduced and accepted, it will detract from the value of Standard English, and consequently, their value as experts. Finally, learners may feel that it is unnecessary for them to study different varieties of English that they do not intend to use and regard as inferior.

Such attitudes have been documented in numerous cases related to bilingual education, for example, in the use of *Tok Pisin* (an English-based Creole language) as a medium of instruction in Papua New Guinea (Romaine, 2000), or in the use of Ebonics in the Oakland California School District (Fasold, 2001, Murray, 2001). While not strictly parallel, these cases illustrate the point that learners (and their parents) have strong opinions about which language varieties are acceptable and prestigious, and which they consider unacceptable or lacking in prestige, and therefore do not want to use nor have their children use—even if using such a variety would facilitate learning. In the case of JE, there is indirect evidence of a pejorative attitude in cartoons that occasionally appear in Japanese periodicals that are based on the misuse or mispronunciation of English words by Japanese. In light of these negative attitudes, one cannot expect that JE will be readily accepted. If there is to be greater acceptance, it can come only through action on the part of teachers, researchers and materials writers.

Conclusion

The WE approach provides a useful perspective for analyzing several important issues concerning the teaching and use of English in Japan. In Japan, as in other places where English is used as a foreign language, one key issue concerns the variety to be used for pedagogical purposes. Usually Standard English is used as a model and other varieties receive only scant consideration. In contrast to this, the WE approach stresses that rather than selecting a single variety as the one to emulate and teach, the important thing is to expose learners to as many varieties as possible so as to prepare them to encounter English as it is actually used in the world. Besides making students aware of the existence of other varieties, this has a beneficial effect on their attitudes: It helps Japanese speakers of English to view speakers of non-Inner Circle varieties in a more positive light, and it could enable them to regard themselves as not as speakers of “broken English,” but as speakers of what should be a recognized variety of English.

Another pedagogically related issue is the native speaker/non-native speaker distinction. Whether or not a teacher is a native speaker tends to be an important factor in selecting teachers, but on examination this distinction seems spurious and irrelevant. In advocating greater acceptance of EFL teachers who are not speakers of Inner Circle varieties, one can anticipate resistance from those who feel that the quality of teach-

ing and learning will be endangered, but it is well to bear in mind that diversity and quality are not incompatible, and that quality itself may be diminished if diversity is not accepted. While a teacher's experience and academic qualifications should remain primary considerations in hiring decisions, being a speaker of an Outer Circle variety need not be seen as a detriment to one's suitability as an EFL teacher. Indeed, it can be viewed as a merit, for it helps to make students more aware of a vital aspect of English: its diversity.

The WE approach is concerned with describing the many varieties of English and the roles they have in the contexts in which they are used. From this perspective, one might question the status of JE and whether it should be considered an independent variety of English. In the preceding section, the findings of discourse level studies of the English used by Japanese L1 speakers were presented, and these findings, taken together, were found to furnish stronger support for JE as an independent variety of English than has previously been offered. Considering JE as an independent variety does not, of course, mean that it is a variety in the same sense as the Inner Circle varieties, or the institutionalized varieties from the Outer Circle. JE is, and will probably long remain, a performance variety. The functions for which it is used are limited, though in the future it may gradually come to be used for an expanded range of functions.

Continuing research on various aspects of Japanese speakers' use of English will enable a more complete description of JE to be made. This will put EFL teachers in a better position to distinguish between errors on the one hand, and JE usages on the other. Teachers can then point out learners' non-standard usages, but where these reflect JE, they may choose to deal with them not as errors to be corrected, but as alternative usages. This would entail teachers discussing with learners the contexts in which a JE usage would be more appropriate (as when interacting with another JE speaker) and the contexts in which it would be less appropriate (for example, in formal writing, or in speaking with an interlocutor who is unfamiliar with JE). Such an approach is very much in line with the current TEFL trend toward communicative competence, for it is concerned with training learners to adjust their use of language (in this case, their choice of variety) to make it appropriate to the context. In this way, the WE approach offers a valuable contribution to the practice of English teaching in Japan.

Acknowledgements

I received valuable comments from Jean D'souza and Satoshi Yamazaki, as well as from the editor and two anonymous reviewers who read an earlier version of this paper.

Phillip R. Morrow is a professor in the Faculty of Foreign Studies at Nagoya Gakuin University where he teaches courses in linguistics and EFL. His research interests are in the areas of World Englishes, discourse analysis, and corpus linguistics.

Note

1. Details of this project are available on the internet at:
<http://leo.meikai.ac.jp/~tono/>.

Received May 12, 2003

Accepted September 18, 2003

References

- Beebe, L. M., Takahashi, T., & Uliss-Weltz, R. (1990). Pragmatic transfer in ESL refusals. In R. C. Scarcella, E. S. Andersen, & S. D. Krashen (Eds.), *Developing communicative competence in a second language* (pp. 55-73). New York: Newbury House.
- Brown, H. D. (1987). *Principles of language learning and teaching* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall Regents.
- Brown, H. D. (2000). *Principles of language learning and teaching* (4th ed.). White Plains, NY: Longman.
- Brown, K. (1993). World Englishes in TESOL programs: An infusion model of curricular innovation. *World Englishes*, 12, 59-73.
- Celce-Murcia, M. (Ed.) (1991). *Teaching English as a second or foreign language*. New York: Newbury House.
- Chaudron, C., & Parker, K. (1990) Discourse markedness and structural markedness: The acquisition of English noun phrases. *Studies in second language acquisition*, 12, 43-64.
- Cook, V. (2001). *Second language learning and language teaching* (3rd ed.). London: Arnold.
- Crystal, D. (1997). *English as a global language*. Cambridge: Cambridge University Press.
- Davies, A. (1995). Proficiency or the native speaker: What are we trying to achieve in ELT? In G. Cook, & B. Seidhofer (Eds.), *Principle and practice in applied linguistics* (pp. 145-158). London: Oxford University Press.

- Davies, A. (1999). Standard English: Discordant voices. *World Englishes*, 18, 171-186.
- Doff, A. (Ed.) (1993). *Cambridge skills for fluency*. Cambridge: Cambridge University Press.
- D'souza, J. (1999). Afterword. *World Englishes*, 18, 271-274.
- D'souza, J. (2001). Indian English and Singapore English: Creativity contrasted. In E. Thumboo (Ed.), *The three circles of English* (pp. 3-17). Singapore: UniPress.
- Fasold, R. W. (2001). Ebonics need not be English. In J. E. Alatis, & A. H. Tan (Eds.), *Georgetown University roundtable on languages and linguistics 1999*. (pp. 262-280). Washington, D C: Georgetown University Press.
- Ferguson, C. (1992). Foreword to the first edition. In B. B. Kachru (Ed.), *The other tongue: English across cultures* (pp. vii-xi). Urbana, IL: University of Illinois Press.
- Harmer, J. (2001). *The practice of English language teaching* (3rd ed.). Harlow: Longman.
- Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.
- Jenkins, J. (2003). *World Englishes: A resource book for students*. London: Routledge.
- Kachru, B. B. (1983). *The Indianization of English: The English language in India*. New Dehli: Oxford University Press.
- Kachru, B. B. (1985). Standards, codification, and sociolinguistic realism: The English language in the outer circle. In R. Quirk, & H. G. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11-30). Cambridge: Cambridge University Press.
- Kachru, B. B. (1986). *The alchemy of English: The spread, functions, and models of non-native Englishes*. Oxford: Pergamon Press.
- Kachru, Y. (2001). Discourse competence in world Englishes. In E. Thumboo (Ed.), *The three circles of English* (pp. 341-355). Singapore: UniPress.
- Kay, G. (1995). English loanwords in Japanese. *World Englishes*, 14, 67-76.
- Koike, I., & Tanaka, H. (1995). English in foreign language education policy in Japan: Toward the twenty-first century. *World Englishes*, 14, 13-26.
- Kubota, R. (1998). Ideologies of English in Japan. *World Englishes*, 17 (3), 295-306.
- Long, M., & Richards, J. (Eds.) (1987). *Methodology in TESOL: A book of readings*. New York: Newbury House.
- Loveday, L. (1981). Pitch, politeness, and sexual role: A exploratory investigation into the pitch correlates of English and Japanese politeness formulae. *Language and Speech*, 24, 71-89.
- Maeda, N. (1989). *Polite expressions of refusal in English: Native versus non-native speaker usages*. Unpublished MA thesis. Arizona State University.

- Maynard, S. K. (1997). Analyzing interactional management in native/non-native English conversation: A case of listener response. *IRAL*, 35(1), 37-60.
- McArthur, T. (Ed.) (1996). *The concise Oxford companion to the English language*. Oxford: Oxford University Press.
- McCrum, R., Cran, W., & MacNeil, R. (1986). *The story of English* (2nd. ed.). London: Penguin.
- McCrum, R. (Series writer), Cran, W. (Producer), MacNeil, R. (Co-writer), & Pett, J. (Director). (1986). *The story of English* [videorecording]. New York: BBC-TV co-production with MacNeil-Lehrer Productions in association with WNET.
- Medgyes, P. (1994). *The non-native teacher*. Oxford: Phoenix.
- Morrow, P. R. (1987). The users and uses of English in Japan. *World Englishes*, 6, 49-62.
- Murata, K. (1995). Repetitions: A cross-cultural study. *World Englishes*, 14(3), 343-356.
- Murray, D. E. (2001). Whose "standard"? What the Ebonics debate tells us about language, power, and pedagogy. In J. E. Alatis & A. H. Tan (Eds.), *Georgetown University roundtable on languages and linguistics 1999* (pp. 281-291). Washington, DC: Georgetown University Press.
- Nayar, P. B. (1997). Variants and varieties of English: Dialectology or linguistic politics? In H. Lindquist, S. Klintborg, M. Levin, & M. Estling (Eds.), *The major varieties of English* (pp. 283-289). Växjö, Sweden: Växjö University.
- Quirk, R. (1995). *Grammatical and lexical variance in English*. Longman: London.
- Romaine, S. (2000). *Language in society* (2nd ed.). Oxford: Oxford University Press.
- Sasaki, M. (1997). Topic continuity in Japanese-English interlanguage. *IRAL*, 35(1), 1-21.
- Shim, R. J. (1999). Codified Korean English: Process, characteristics and consequence. *World Englishes*, 18(2), 247-258.
- Song, J. J. (1998). English in South Korea revisited via Martin Jonghak Baik (1992, 1994), and Rosa Jinyoung Shim (1994). *World Englishes*, 17(2), 263-271.
- Stanlaw, J. (1982). English in Japanese communicative strategies. In B. B. Kachru (Ed.), *The other tongue: English across cultures* (pp. 168-197). Urbana, IL: University of Illinois Press.
- Suenobu, M., Yamane, S. & Kanazaki, K. (1997). Information transmission of English by Japanese learners of English. *IRAL*, 35(3), 187-197.
- Takashi, K. (1990). A sociolinguistic analysis of English borrowings in Japanese advertising texts. *World Englishes*, 9(3), 327-341.
- Thumboo, (Ed.). (2001). *The three circles of English*. Singapore: UniPress.
- Ur, P. (2001) *A course in language teaching: Practice and theory*. Cambridge: Cambridge University Press.
- White, S. (1989). Backchannels across cultures. *Language in society*, 18, 59-76.

- Widdowson, H. (1996). The ownership of English. *TESOL Quarterly*, 30, 379-389.
- Yamada, H. (1990). Topic management and turn distribution in business meetings: American vs. Japanese strategies. *Text*, 3, 271-295.
- Yano, Y. (2001). World Englishes in 2001 and beyond. *World Englishes*, 20, 119-131.

Reviews

Task-based Language Learning and Teaching. Rod Ellis. Oxford: Oxford University Press, 2003. x + 387 pp.

Reviewed by

Thomas C. Anderson

Aoyama Gakuin University

Dr. Rod Ellis, a leading figure in Second Language Acquisition research, is able to synthesize large amounts of research and explain it clearly to language educators, as can be seen in his latest work *Task-based Language Learning and Teaching*. Ellis' bibliography, 24 pages long and containing 637 entries spanning five decades, shows the effort he put into this book.

In the preface, he shares three reasons for this work. First, he is a firm believer in language teaching using language as a classroom communication tool in preparation for real world communication. He views tasks as the best way to achieve this. Second, he says that tasks have become a central concept in SLA research, both as instruments for examining L2 acquisition and as something to be investigated themselves. Finally, and most significantly, he believes that task study can help bridge the gap between SLA research and classroom teaching practice.

There are two caveats in the preface. First, the book's contents are limited mostly to psycholinguistic accounts of tasks as these comprise Ellis' area of expertise. Second, the book is not a "how to" recipe book for teachers. It is designed to identify both the advantages and the downside of task-based teaching. Teachers should come away with a number of ideas, which they can try out and adapt to their classroom situation.

The book begins by examining tasks as an SLA research concept, presenting several definitions and describing their critical features. Ellis distinguishes between unfocused tasks and those that are used to elicit a particular linguistic feature or make language itself the task content. Following this, he looks at the classroom use of tasks.

He mentions the difference between task-supported language teaching, in which tasks are "a means by which learners can activate their

existing knowledge of the L2 by developing fluency,” and task-based language teaching, in which tasks are “the basis for an entire language curriculum” (p. 30).

The next three chapters look at unfocused tasks. First, listening to comprehend and listening to learn are examined. Following this, tasks involving interactions between learners or between learners and teachers/researchers are described in terms of negotiation of meaning, communication strategies, and interaction. The last chapter in this section focuses on the content produced by learners and how it relates to language acquisition. Ellis concludes by considering the operationalization of production in various studies to see how production relates to task design and implementation.

In chapter 5 Ellis turns to focused tasks. He describes their psycholinguistic basis and three principal designs: structure-based production tasks, comprehension tasks, and consciousness-raising tasks. Finally, he looks at implicit and explicit techniques used in implementing focused tasks.

In chapter 6, Ellis switches from the “black box” metaphor that usually is used to inform mainstream SLA, to an intra/interactional focus of sociocultural SLA. Learning, in this view, is dialogically based. Speech, both to others and oneself, plays a crucial role. Orientation and student attitudes, as well as scaffolding between the teacher/researcher and learners or among learners themselves, are also important in developing awareness/production of new L2 forms and assisting in acquisition. Ellis ends the chapter by calling for a pluralistic approach to task-based language learning.

The next chapter looks at the practical concerns that arise with task-based program innovation: development of a task-based course curriculum, task-based teaching methodology, and task-based assessment issues. The book concludes with a chapter on task-based pedagogy evaluation.

Reviewing Ellis’ three purposes (emphasizing oral communication, explaining the nature and importance of tasks in SLA research, and bridging the gap between SLA research and teachers), it is clear that he achieves these goals. The reader will gain a better understanding of how tasks relate to SLA, the difference between unfocused processing and interaction tasks and those that are focused, the relationship between tasks and sociocultural SLA, and the implementation of tasks at the course level including curriculum planning and development, classroom use, and class/program assessment and evaluation.

With this new understanding of the concepts and issues involved in task design and use (aided by the concisely written glossary and extensive bibliography), educators may become excited as they become aware of the action research possibilities in their classroom. This is perhaps the most important reason for using tasks.

Methodology in Language Teaching: An Anthology of Current Practice. J. Richards & W. Renandya (Eds.). Cambridge: Cambridge University Press, 2002. vii + 422 pp.

Reviewed by

Roger Kenworthy

Ohio University, Hong Kong

Recently, a major paradigm shift has occurred with the Communicative Approach emerging as the principal tenet for present day second language instruction and learning. *Methodology in Language Teaching: An Anthology of Current Practice* by editors Richards and Renandya is a comprehensive overview of the Communicative Approach with practical activities tailored to elementary, secondary, and tertiary level classrooms, and a source for pre- and in-service teacher-education programs.

The authors divided the textbook into sixteen sections, each with a thoughtful introduction and an ample supply of references intended to benefit researchers and instructors. Every section is guided by a series of questions. The pre-reading questions draw upon personal beliefs and experiences as a source of reference for the articles, while the post-reading questions both engage personal reflection and challenge the readers to apply a wide variety of activities and methods to their teaching practice.

One of the strengths of the work is the breadth of topics ranging from *Approaches to Teaching, Learning Strategies, and Technologies in the Classroom* to *Professional Development*. However, the main focus of the text is upon the nuts and bolts of the Communicative Approach with a thorough discussion of the teaching of grammar, pronunciation, speaking, listening, vocabulary, reading, and writing. The salient points

for each language skill conveyed by the contributors are summarized, beginning with grammar.

There is an uneasy tension among Communicative Approach grammar theorists and practitioners over the debate of fluency versus accuracy. Within the Communicative Approach, researchers argue that grammatical accuracy is essential within all skills while practitioners focus upon the use of real and practical communication which often means that fluency supersedes grammatical accuracy. Ultimately, it must be the aim of instruction that resolves this contentious issue.

Pronunciation practices have been changing dramatically as the focus has shifted from accurate production of individual sounds to the global features of stress, rhythm, pitch, and intonation. Also, because the purpose of speech varies, speakers need to become fluent and use appropriate language required for successful communication. As such, speaking courses should address the complex and varying needs of language learners.

Until recently, listening was thought to be acquired through repeated exposure with little need for explicit instruction. However, in Chapter 22, Field opines that theory has changed to provide a more focused approach as instructors actively guide learners through the processes of learning and identifying common difficulties in order to make appropriate changes in the classroom. The outcome should present learners with the maximum opportunity for personal involvement and development of their listening skills.

There is renewed interest in vocabulary teaching and learning because it is the core for speaking, listening, reading, and writing. Traditionally, vocabulary instruction often meant presenting new words de-contextualized and having no meaningful relationship with other words. However, a paradigm shift has occurred and words are thought of in a much broader sense that includes the learning of phrases, idioms, and collocations. In this section, Nation provides a number of strategies for learning vocabulary within a network of communicative tasks.

Currently there is considerable debate among language researchers on the place of reading in the second language writing classroom. In Section 12, Grabe argues that extensive reading is crucial to learners as they are put in contact with "good models" of organization, varied vocabulary, and grammar. Also, the Communicative Approach favors the process approach to teaching composition. With four distinct stages in the process (preplanning or pre-drafting, drafting, revising, and editing),

this is a systematic but flexible approach to composition allowing writers to know where they are and what they should focus upon at every step of the process. In this section, Seow maintains that learners' writing competency increases due to the interaction between the process, writer, and task, rather than from the learning of structures.

The strength of this highly recommended text lies within the breadth of subjects covered which enables all levels of instructors to benefit from the authors' skillful balance of theoretical background information and practical exercises. This book defines the state of second language instruction and learning at the beginning of the millennium and is a must for your bookshelf.

Practical English Language Teaching (International Edition). David Nunan (Ed.). New York: McGraw Hill, 2003. xi + 342 pp.

Reviewed by

Paul Hullah

Nagoya University of Business

The stated aim of this comprehensive survey of language teaching methodology is to "update the experienced teacher on current theoretical and practical approaches" and offer the "novice teacher ... step-by-step guidance on the practice of language teaching" (p. vi). It takes on this task by providing, in a set chapter by chapter format, chronologically-framed overviews up to present-day opinion of 15 ELT areas housed in three sections: "Exploring Skills," "Exploring Language," and "Supporting the Learning Process." Each chapter is authored by a different expert in the particular field. Nunan writes two, but then he is the editor and, as anyone familiar with his other 30 TEFL book publications knows, an expert in numerous fields.

Each chapter defines its subject skill (e.g. "Listening," "Pronunciation," "Coursebooks," "Learner Autonomy"), presents the historical background, summarizes research, and expands upon relevant concepts from a four-skills-based perspective. Factors necessary for approaching the skill practically as a teacher are then outlined: examples from teaching

materials are included, with transcripts of actual classroom interchanges illustrating an effective or ineffective pedagogical approach. Chapters close with links to suggested further reading and appropriate web sites.

This book is a valuable resource to have on one's bookshelf, especially for teachers adopting an integrated skills approach. Chapters are consistently lucid and intelligently cross-referenced. Ken Beatty's "CALL" chapter deals reasonably with the rapidly evolving sphere of online language learning, explaining the usefulness of software-based constructivist programs and "blended learning" sessions with teachers in "chartrooms" (*sic*), presumably *chatrooms*, the only typographical slip I found in the whole book. Maggie Sokolik tackles an intimidating subject ("Writing") enthusiastically, advocating a principled eclectic approach and examining evaluative methods. Pleasingly, the style of each individual contributor is allowed to peek through the uniformly precise prose. Thus Kathleen Bailey ("Speaking") manages a witty dig at the woodenness of inauthentic audiolingual texts, while Donna Brinton ("Content-based Instruction") gives personal anecdotal description of her attempts to learn German. Unintentional irony also occurs: EFL teachers in Japan may groan to find that three examples of outdated, "not... useful" discrete-point test-items "from the 1950s to about the mid-1970s," cited by Geoff Brindley ("Classroom-based Assessment") are precisely the kind of questions which still dominate the English section of present-day Japanese university entrance exams.

Not so dependably well-conceived as the main text, however, are the "Reflection" and "Action" boxes punctuating each chapter, "inviting readers to reflect on issues, principles, and techniques... [and] apply the ideas through action-oriented tasks" (p. vii). Though most of these are provocative and useful, some appear hastily or lazily formulated, for example the "Reflection" sections of the "Vocabulary" and "Content-based Instruction" chapters. Some of these "inserts" pose direct, quantifiably answerable questions, but no answers are provided in subsequent text. The rhetorical or open-ended nature of such tasks ("Name as many countries as you can where English is the dominant language..." p.112) is frustrating: How can a reader gauge comprehension of a principle when questions testing that comprehension are left unanswered? Perhaps each author designed asides for his or her section? If this is the case, Nunan, as editor, might have been better employed composing all the boxes himself. Tellingly, his own "Grammar" chapter inserts are excellent, each followed by an explanatory "Commentary" absent from some other chapters.

However, that is a minor quibble. This is an excellent book. Few assumptions are made of a reader's background knowledge, but passages which veteran EFL instructors might find old hat are presented in a lively, concise manner that never patronizes or bores. For most teachers, parts of this well-written book will feel like shaking hands with old friends; other parts will feel like meeting new ones.

Teachers' Narrative Inquiry as Professional Development. Karen E. Johnson & Paula R. Golombek (Eds.). Cambridge: Cambridge University Press, 2002. 212 pp.

Reviewed by

Michael J. Crawford

Hokkaido University of Education, Hakodate

Professional development comes in many shapes and forms in the field of language teaching. Common options include attending workshops and conferences, enrolling in certificate courses, and reading journal articles and books. In many cases, these forms of professional development can be considered part of what Johnson and Golombek describe in their book *Teachers' Narrative Inquiry as Professional Development*, as the transmission model of teacher education. In this model, educational researchers impart specialist knowledge to teachers, who are expected to learn and apply it in their teaching.

For language teachers who have harbored doubts about the efficacy of the transmission model of teacher education, this book will undoubtedly strike a chord. The editors argue that the transmission model marginalizes teachers by making them passive recipients of knowledge, as opposed to active creators of knowledge. As an alternative, Johnson and Golombek propose a model of professional development "conducted by teachers and for teachers" (p. 6). At the heart of their model is storytelling or "narrative inquiry."

Teachers' Narrative Inquiry as Professional Development is a collection of stories told by language teachers. In these stories, teachers describe in detail aspects of their teaching that they are consciously seek-

ing to understand. This method of professional development, the act of relating a story about one's teaching, turns teachers into active creators of knowledge about themselves, their educational settings, learners, and teaching practices.

The book is divided into fourteen chapters. In the first chapter, the editors introduce narrative inquiry and note the limitations of the transmission model of professional development. They also acknowledge the limitations of narrative inquiry. The rest of the book is divided into four parts. Part I focuses on instructional practices, Part II on learners, Part III on teachers, and Part IV on professional collaborations. In this review, I will describe one chapter each from two of the four parts.

In Part I teachers write about aspects of their teaching with which they are dissatisfied. In chapter 3, Lynne Doherty Herndon expresses concern that her love of literature causes her to dominate the class. She experiments with three different approaches to teaching reading that aim to create a more student-centered atmosphere. This experimentation leads to changes in her instructional practices thus creating a more satisfying classroom experience for both her and her students.

The chapters in Part IV are written by teachers who seek the help of colleagues to solve a problem with their teaching. In chapter 13, Michael Boshell asks a colleague to help him understand why he is unable to get quiet students to participate in class. He learns that his teaching style denies these students space, both in the physical sense, because he watches over them too closely, and in the psychological sense, because he attempts to control too much what they say. The process of examining this problem with a colleague allows him to make changes in his teaching style.

One of the strong points of *Teachers' Narrative Inquiry as Professional Development* is the variety of stories. In addition, the stories are thought-provoking, and at times humorous and heartwarming. Although the stories may resemble reports on action research, they differ in that they give readers a better chance to look at the big picture. The effects that teachers' personalities, interpersonal relationships, philosophies, and experiences have on their teaching are revealed in a way that is not usually found in action research.

The highly personal nature of the stories in *Teachers' Narrative Inquiry as Professional Development* can be seen as both a strength and a weakness of the book. Readers looking for generalizable data that can be used to support specific approaches to language teaching or teacher

training will be disappointed. To offer this as a criticism, however, would be unfair as this is clearly not the purpose of the book. Rather, the editors have sought to make teachers' thoughts about their craft open to public view, in the hope that other teachers will be stimulated to reflect on their own teaching practices. On this score, the editors are successful. Readers will find themselves thinking about their teaching practices and may very well want to tell their own stories.

Controversies in Applied Linguistics. Barbara Seidlhofer (Ed.). Oxford: Oxford University Press, 2003. xiv + 346 pp.

Reviewed by
Scott Petersen
Seijoh University

Applied linguistics draws on various fields. The three main areas are linguistics, psychology (or psycholinguistics), and sociolinguistics. The controversies that Seidlhofer has chosen come from the latter. If you want to read about psychological/psycholinguistic controversies, then look elsewhere. This book focuses exclusively on controversies concerning language and society. Except for one section, it does not address everyday concerns of language teachers.

Seidlhofer has selected sets of already-published articles that address a particular controversy. Each set begins with an article that raises an issue of contention, followed by articles that respond to it. Usually, the first article addresses problems in the work of contemporary scholars, who in subsequent articles refute or at least indicate what they see as misunderstandings. Often the original author then writes a refutation to the refutation. By following several controversies through the stages of argument, counter-argument, and counter-counter argument, the editor hopes that readers will gain an appreciation of how controversy in academia proceeds. Consequently, people interested in theoretical discussion will most enjoy this book. In addition to choosing the articles, Seidlhofer also sandwiches each controversy with an introduction, afterthoughts and a list of material for further study. At the end of the book, she provides general study questions.

Seidlhofer has selected nine controversies which she groups into five sections: World Englishes, corpus linguistics, critical discourse analysis, second language acquisition studies, and the nature of applied linguistics. Since the second section is the only one that directly addresses language teaching, I will briefly comment on the other sections before concentrating on the second.

The controversy surrounding World Englishes revolves around who owns English. Should all people in the world be forced to learn the variants of English in the English-speaking countries? Critical discourse analysis injects sociological concepts into linguistics, thus politicizing it considerably. This section is not really a prime example of academic discussion since the discussants mostly talk past each other rather than to each other. The lead article in the fourth section attempts to reorient the young field of second language acquisition studies so that it takes more notice of the social aspects of learning and less of psychological aspects. The last section raises questions about the exact nature of applied linguistics.

The second section contains the most useful material for language teachers: corpus linguistics and language teaching. At first glance, it would seem that the topic would allow for little controversy. However, while some researchers investigating corpora claim that new, exciting discoveries about language need to be incorporated into teaching materials since ignoring these ideas does learners a disservice, dissenting articles in this book point out flaws in the argument.

The main objection concerns pedagogy. The descriptions of language found in corpus studies are quite complex. How is this complexity to be broken down into presentable and learnable chunks? Maybe learners have no need or desire to emulate, for example, a native speaker of British English in pronunciation or culture-specific turns of phrase. A second objection is that even as descriptions of language, corpus studies fall short. For example, Guy Cook, in text 14 of the book, points out that a corpus of spoken language will show that the word *bet* most often appears in the phrase "I bet." Therefore, the most frequently occurring meaning is closer to *suppose*. However, for the majority of speakers, the core meaning of the word *bet* remains *wager*, and this is in fact the first meaning of the word in the corpus-based COBUILD dictionary. Furthermore, corpora record facts about behavior, not the organization of knowledge in the mind. The last argument turns the tables on linguists whose objection to formal, Chomskian linguistics is that it focuses only on language in the mind and ignores language as behavior. Seidlhofer

ends this section with a list of resources for pursuing the controversy. Even though hardliners may overstate their case, corpus linguistics does have application in language teaching. Indeed, the third issue of *TESOL Quarterly* for 2003 deals exclusively with corpus linguistics.

Seidlhofer has done a fine job of introducing these controversies and providing references for further study. Those interested in the interplay between language and society will enjoy the book. Teachers and material writers will gain much food for thought in the corpus linguistics controversy.

Applied Linguistics in Language Education. Steven McDonough.
London: Arnold Publishers. 2002. 178 pp.

Reviewed by
Scott Bronner
Sophia University

Steven McDonough does an excellent job of reviewing the controversies and current questions in the field of Applied Linguistics. His style is refreshingly honest and practical, using real classroom examples as a framework to discuss key issues. For those who have taught English and who have some background in Applied Linguistics research, this text will be stimulating and helpful in consolidating ideas useful to teaching and action research. However, anyone without such a background or explanatory resource text, may find disturbing the lack of definition of key terms and even acronyms. Therefore I would recommend *Applied Linguistics in Language Education* for those familiar with the field and teachers desiring to consolidate and add to their knowledge of Applied Linguistics.

This text certainly approaches Applied Linguistics from the “applied” more than the “linguistic” perspective, with the connection of research to teaching being one of the key issues covered. McDonough is also unabashedly critical of various teaching methods that have come and gone and of the assumption that a relatively narrow methodology can be found that will be more effective than an eclectic approach guided by experience and practical research. Thus the text glosses over methodologies such as the Audiolingual Method and Natural Approach, while

other methods like “the Silent Way” are not even mentioned. Instead, McDonough focuses on issues relating clearly to current classroom teaching and highlights how research may be conducted in a way that will have either immediate or long-term applicability.

Major sections in the book are “What is applied linguistics?” “Language, linguistics, and teaching,” “Language learning,” and “Applied linguistics and the teaching profession.” Subsections discuss a wide range of topics, including how the field interacts with areas such as language education and linguistics, textual issues, interlanguage studies, the concept of method, teacher development, and assessment and evaluation.

Overall, I found the book to be interesting with a substantial amount of thought-provoking material. Particularly helpful were the examples of various types of data used in Applied Linguistics, from a lesson transcript demonstrating classroom language issues to a learner’s protocol in which someone writes down her thoughts while completing a task in a second language. These examples are presented in the beginning and referred to throughout the text. Accordingly, McDonough chooses to approach this text through the lens of various kinds of textual data instead of presenting a dry account of research methods or a traditional history of language-teaching methodologies and movements.

My main concern is that without a glossary of terms, even mention of important innovations lack clarity, such as the applicability of discourse analysis and corpus linguistics to language teaching. Hence it may be confusing when McDonough goes into related issues, for example, whether turn-taking or the cohesive device of anaphoric reference via pronouns should be treated as part of language structures or language skills (p. 43). This lack of explanation is apparent throughout the book, including the use of undefined acronyms, such as “EST” (p. 49).

I recommend *Applied Linguistics in Language Education* to those looking for a concise account of the issues in our field” and who are not disturbed by a book that raises more questions than it answers. However, since it assumes a certain amount of background information, the text could frustrate those unfamiliar with Applied Linguistics. Finally, I wholeheartedly agree with the book’s consistent plea for a more interdisciplinary approach that would better connect Applied Linguistics with research conducted in education, psychology, and other fields.

The Changing Face of CALL: A Japanese Perspective. Paul E. Lewis (Ed.). The Netherlands: Swets & Zeitlinger, 2002. pp. xiii + 274.

Reviewed by

Justin Charlebois

Nagoya Bunri University

Although Japan has maintained a reputation as an innovator in the development of technology, most of its educational technology is not exported. Furthermore, much of the literature about Computer Assisted Language Learning (CALL) in Japan is written in Japanese, and thus not readily accessible to other parts of the world. The current volume is a collection of papers gathered from the JALT CALL SIG 2001 conference that was held at Kanto Gakuen University in Gunma Prefecture.

The book is divided into three main sections, with a total of 17 chapters. The first section, "Theories, Models, and Paradigms," has chapters about applying hypertext concepts to language acquisition, the distinction between CALL and Artificial Intelligence (AI), and even Multiple Intelligence (MI) Theory and CALL. The second section, "CALL Resources," has sections on the potential for computerized bilingual dictionaries to enhance vocabulary learning and security on the Internet. The security section provides a lot of useful, as well as slightly alarming, information on viruses. Finally, some of the topics in the third section, "The CALL Classroom and Beyond," include: network-based language teaching, keypal exchanges, teaching cultural awareness through Internet writing projects, and mobile learning.

One of this book's strengths is that it includes a variety of topics about CALL written by different authors. Additionally, the book does not require extensive technical knowledge to be understood. Since the selection of topics is wide ranging, it should appeal to both those with a strong interest in computers and those who are not yet comfortable with the medium but want to become educated about it.

The book also discusses Japan's slow adoption of computers into education. One reason cited, primarily cultural in nature, is that there has always been a higher value placed on penmanship and calligraphy than in the West. Another factor for technology's peripheral role is that language teachers are unsure how to integrate it into the classroom, most likely due to lack of training.

Although there is a wide range of topics discussed in the book, I think the editor effectively organized the contributions around three centralized themes. This book will be especially useful for anyone who wants to learn how to integrate computers into the language classroom, and for those who want some background knowledge about the work that has already been carried out. While there are many books on the market about CALL, the most useful aspect of this book is that it specifically addresses the Japanese educational context.

Pragmatics. George Yule. Oxford: Oxford University Press, 1996. xiv + 138 pp.

Reviewed by

Donna Tatsuki

Kobe City University of Foreign Studies

This book is one of the oldest contributions to the *Oxford Introductions to Language Study* series. The series was “designed to provide [a] large scale view of different areas of language study” (p. xi) which makes it an excellent preliminary text series for a general readership interested in language and for the serious student of linguistics. Pragmatics had long been characterized as the linguistic “wastebasket” (p. 6) by grammarians who could not reconcile examples of actual language use with the tidy syntactic and semantic categories they had constructed. This changed in the mid-1970’s when linguists such as Austin, Searle, and Wilson, as well as conversation analysts such as Sacks, Schegloff, and Jefferson started sifting through and trying to make sense of the contents of this overflowing recycle bin. Their pioneering efforts resulted in the vibrant field of pragmatics that we know today. It was fitting that one of the earliest and most prolific contributors to this field, George Yule, was invited to contribute a volume on pragmatics in the Oxford series.

Like the other books in the series, *Pragmatics* has four sections: survey, readings, references, and glossary. The survey, which takes up the majority of pages, offers chapters such as “Politeness and Interaction” and “Conversation and Preference Structure.” In this section, George Yule has masterfully organized pragmatics into eight areas of focus beginning

at the word level, “Deixis” (Latin for “pointing”), and gradually enlarging the scope with each chapter until it culminates in the final chapter, “Culture and Discourse.” The important technical terms preferred by specialists in the field are printed in bold face type along with individual entries in the *Glossary* (Section 4). The most difficult chapter to read, in my view, was the one entitled “Presupposition and Entailment” with its use of symbols and logic equations. To be fair, this could be more my problem than that of the text.

Section 2: Readings features a collection of short passages that illustrate and expand key points raised in the *Survey* section. The best parts of this section are the follow-up questions that Yule poses to encourage readers to go beyond the reading by thinking of their own examples and even proposing alternative explanations for phenomena. For example, after the chapter 1 supplementary reading passage by Georgia Green, Yule challenges the reader as follows:

From this description it seems as if every act in life is part of pragmatics. Do you think that pragmatics is the study of all actions, or should it be limited to only certain actions? What kind of limitations would you propose? (p. 91)

Yule has also ensured that this book is relevant to readers from a wide variety of backgrounds by pointing out Anglo-centric notions and encouraging readers to reframe the arguments in their own contexts. For instance, one of the follow-up questions for the chapter 5 supplementary reading passage by Paul Grice is:

Grice emphasizes the word “reasonable” as he describes his consideration of the cooperative principle and his maxims as a kind of contract. Would the cooperative principle, the maxims, and the three features listed here be treated as “reasonable” in all societies and cultures? (p. 101)

Section 3: References offers graded references that I have found particularly helpful for directing my students to appropriate and accessible articles on these topics. However, there are a couple of references (dissertations and conference proceedings) that would be difficult to obtain so one might wonder if their inclusion is really useful. Nevertheless, I would heartily recommend this text to anyone who would like a compact, concise introduction to pragmatics.

The Bilingual Family: A Handbook for Parents (Second Edition).
Edith Harding-Esch & Philip Riley. Cambridge: Cambridge University
Press, 2003. xvi + 190 pp.

Reviewed by

Jack Massalski

Tokyo Woman's Christian University

Being trilingual, or raising a family where neither of the parents uses their native language in the home can seem confusing to some. However to others, like myself (a Polish Canadian living in Japan), cultivating bilingualism at home seems natural and appropriate. At the same time, being a professional linguist who is familiar with current research on bilingualism does not completely prevent one from occasional moments of uneasiness regarding personal linguistic choices, and pondering whether cultivating a bilingual family is worth the effort. Wanting to know more about how others have coped with the challenges of bilingualism in the family, I examined the book in question and was not disappointed.

I strongly recommend *The Bilingual Family: A Handbook for Parents* by Edith Harding-Esch and Philip Riley as an excellent resource for parents who wish or need to raise their children as bilinguals. The book's anecdotal and optimistic tone makes it an easy and enjoyable read for those interested in the nature of bilingualism. As the title implies, the book serves primarily as a parenting handbook and thus a background in linguistics is not necessarily required of the reader.

The second and updated edition of the book reviewed here consists of three main sections. Part 1, "A survey of the issues," contains a summary of research conducted on bilingualism and the development of the bilingual child, as well as a discussion of various factors that parents should take into consideration when making decisions regarding the child's bilingualism. Parents will be relieved to learn that mixing languages is a natural part of "sorting it out" by the child and that extensive language correction is not necessary. In my opinion, the first part of the book alone is worth the price of the volume.

In the following section, the authors present a series of case studies that expose the reader to various linguistic arrangements in the family, the choices that each family made, and how the choices affected their children. From reading the studies the reader will get the overall posi-

tive impression of naturalness with which children manage to acquire languages. In this updated edition, some studies also include follow-up interviews with children to show how important being bilingual is to their current identity. Interviewees highly value their bilingualism, regarding it as an enrichment and express gratitude to their parents for providing them with a bilingual environment at home.

The final section of the book is a glossary of terms, topics and notions that may be useful to parents seeking advice regarding bilingualism. In this section, the reader will find discussions of topics such as accent, baby talk, code-switching, code-mixing, swearing, and writing, as well as practical tips and various suggestions meant to assist and encourage the parent of a bilingual child.

My only criticism of the book is that it primarily deals with European languages, mainly French, English, and German. In future editions I would certainly like to see included case studies of families facing the challenges of raising bilingual children of languages strikingly different from each other in terms of grammar, lexicon, and most importantly, the culture they operate in. My assumption is that it would normally take longer for a child acquiring distant languages such as English and Japanese to sort them out than for a bilingual child of English and French under similar circumstances. It would also be interesting to observe how children cope with the cultural differences associated with the two languages (a topic the book in question unfortunately does not even touch upon) and how they become bicultural as well as bilingual. Obviously more research is needed in this area.

Putting minor criticisms aside, *The Bilingual Family* is an excellent resource that is interesting, informative and fun to read. If you are a parent wanting to raise your child bilingually, this book will certainly be very useful to you as a reference. On the other hand, if you are debating whether a bilingual family is what you really want, this book will answer many of your questions. Finally, if like me, you are already a parent of a bilingual child and need reassurance or support in your struggle, this book is definitely for you. Buy it. You will not be disappointed.

The Japanese Mind: Understanding Contemporary Japanese Culture.
 Roger J. Davies & Osamu Ikeno (Eds). Tokyo: Tuttle Publishing, 2002.
 vii + 270 pp.

Reviewed by

Jonathan M. W. Rankine
 strategic.comMUNICATION

At 270 pages, this is a slim collection of essays on “key concepts in Japanese culture” (p. 1). Intended as a text, each of the 28 essays is followed by discussion questions which are separated into two groups: one for Japanese students of EFL and the other for foreign students of Japanese Studies. Furthermore, the co-editors intended that through clarity, well-documented research, and demonstrated field-testing, the text would also appeal to the general reader.

Unfortunately, this text fails on almost all accounts. Written by Japanese undergraduate seniors, the explanations are simplistic, superficial, and inconsistent. The first essay on the purportedly unique-to-Japan *chinmoku* (silence) is an illustration. It is used during times of thoughtfulness, hesitation, restraint, conflict avoidance, defiance, and indifference, in public and in private (pp. 53-55). This “unique” Japanese cultural trait has been defined so broadly as to become meaningless, since it covers almost every moment of silence one could experience anywhere.

The superficiality of the research is reflected in the use of E. Reichauer’s (1990) comments originally made in 1977 on the contemporary status of marriage in Japan: “Japanese women are often said to have difficulty in socializing freely... However, women seem willing to play their own roles in maintaining the household as good wives and mothers” (p. 67). One wonders how “freely” socializing women or “good wives and mothers” who are unhappy with their roles and divorce their husbands fit into these nearly thirty-year-old arguments. There is also the incorrect statement that White Day is only found in Japan (p. 98). It is also found in South Korea. Furthermore, there is an inconsistent level of analysis. *Honne* and *tatemaie* (private versus public persona) receive only two pages of text, but *soshiki* (funerals) receives 14 pages, even though the latter is high on detail and low on analysis.

However, this text's greatest weakness lies in the editing, for, as the editors admit, the essays are patchworks of many papers on the same

or similar topics, which is why no single essay is credited to any one author. The results are frequent jumps in argumentation and awkward or altogether puzzling insertions within the essays, as well as much overlap and repetition among the essays. For example, the concept of *amae* is defined twice and explained multiple times (pp. 17-19, 67, 103-104). The concept of *vertical society* is defined three times (pp. 10-11, 144, 187-188). Both *honne* and *tatemaie* (pp. 104-105, 115-116, 195) as well as *ie* (pp. 61-62, 119-124, 217-218) are defined three times. Many other concepts are similarly over-defined. There are also basic grammatical and sentence structure errors, including run-on sentences, capitalization, and verb-agreement problems. It is surprising that this book was edited by two professors and has gone through Tuttle's editing process.

The book's basic premise is to explain and create discussion on contemporary Japanese culture. However, it is centered on a historical Japan that not only has changed, but also is changing in many of the areas covered. Not to be found are discussions on contemporary Japanese cultural traits exemplified by *enjokosai* (teenage prostitution), *furiita* (young, part-time workers with little hope or belief in the future), or *tomodachi-oyako* (an unhealthy parent-child friendship deficient in minimal socialising functions that are usually derived from parental hierarchy). From these (admittedly negative) contemporary Japanese cultural traits there is much to be mined, such as the fixation on youth, with the inherent fetishising of school girls and pressure on older women and mothers to be young and girlish, and the effects of 10 years of economic decline on a disenfranchised youth.

This text presents concepts that fit in with the tea garden and mossy stone view of Japan, while in reality, Japanese culture is a vibrant and dense culture in flux, equally as modern as any other. Unfortunately, poor research, writing, and editing misrepresent traditional cultural traits while neglecting contemporary ones. For sociological analyses of Japan, the reader should stick to monographs put out by trained sociologists. Perhaps the flip side of that is linguists should tread carefully in areas that are not their expertise.

References

- Reischauer, E. (1990). *The Japanese today: Change and continuity*. (M. Fukushima, trans.). Tokyo: Bungei Shunju. (Original work published 1977.)

Information for Contributors

All submissions must conform to *JALT Journal* Editorial Policy and Guidelines.

Editorial Policy

JALT Journal, the refereed research journal of the Japan Association for Language Teaching (Zenkoku Gogaku Kyoiku Gakkai), invites practical and theoretical articles and research reports on second and foreign language teaching and learning in Japanese and Asian contexts. Submissions from other international contexts are accepted if applicable to language teaching in Japan. Areas of particular interest are:

1. Curriculum design and teaching methods
2. Classroom-centered research
3. Cross-cultural studies
4. Testing and evaluation
5. Teacher training
6. Language learning and acquisition
7. Overviews of research and practice in related fields

The editors encourage submissions in five categories: (1) full-length articles, (2) short research reports (*Research Forum*), (3) essays on language education or reports of pedagogical techniques which are framed in theory and supported by descriptive or empirical data (*Perspectives*), (4) book and media reviews (*Reviews*), and (5) comments on previously published *JALT Journal* articles (*Point to Point*). Articles should be written for a general audience of language educators; therefore statistical techniques and specialized terms must be clearly explained.

Guidelines

Style

The *JALT Journal* follows the *Publication Manual of the American Psychological Association*, 5th edition (available from APA Order Department, P.O. Box 2710, Hyattsville, MD 20784, USA; by e-mail: <order@apa.org>; from the website: <<http://www.apa.org/books/ordering.html>>). Consult recent copies of *JALT Journal* or *TESOL Quarterly* for examples of documentation and references. A downloadable copy of the *JALT Journal* style sheet is also available on our website at <<http://www.jalt.org/jj/>>.

Format

Full-length articles must not be more than 20 pages in length (6,000 words), including references, notes, tables, and figures. *Research Forum* submissions should be not more than 10 pages in length. *Perspectives* submissions should be not more than 15 pages in length. *Point to Point* comments on previously published articles should not be more than 675 words in length, and *Reviews* should generally range from 500 to 750 words. All submissions must be typed and double-spaced on A4 or 8.5"x11" paper. The author's name and identifying references should appear only on the cover sheet. Authors are responsible for the accuracy of references and reference citations.

Materials to be submitted

1. Three (3) copies of the manuscript, with no reference to the author. Do not use running heads.
2. Cover sheet with the title and the author name(s).
3. Contact information, including one author's full address and, where available, a fax number and e-mail address.
4. Abstract (no more than 150 words).
5. Japanese translation of the title and abstract, if possible (less than 400ji).
6. Biographical sketch(es) (no more than 25 words each).
7. Authors of **accepted** manuscripts must supply camera-ready copies of any diagrams or figures and a disk copy of the manuscript (RTF or ASCII).

Evaluation Procedures

Submissions will be acknowledged within one month of their receipt. All manuscripts are first reviewed by the editorial board to insure they comply with *JALT Journal* Guidelines. Those considered for publication are subject to blind review by at least two readers, with special attention given to: (1) compliance with *JALT Journal* Editorial Policy, (2) the significance and originality of the submission, and (3) the use of appropriate research design and methodology. Evaluation is usually completed within three months. The first author of a published article will receive one copy of the *Journal* and 10 off-prints with the option to order more at a set rate (contact JALT Central Office for price details). Review authors receive one copy of the *Journal*.

Restrictions

Papers submitted to *JALT Journal* must not have been previously published, nor should they be under consideration for publication elsewhere. *JALT Journal* has First World Publication Rights, as defined by International Copyright Conventions, for all manuscripts published. We regret that manuscripts or computer disks cannot be returned. In the interest of facilitating clarity, the editors reserve the right to make editorial changes to accepted manuscripts.

Full-Length Submissions, Research Forum, and Point to Point Submissions

Please send submissions in these categories or general inquiries to:

Donna Tatsuki, *JALT Journal* Editor
Kobe City University of Foreign Studies
9-1 Gakuen Higashimachi, Nishi-ku, Kobe 651-2187, Japan

Perspectives

Please send submissions in this category to:

Steve Cornwell, *JALT Journal* Associate Editor
Osaka Jogakuin Junior College
2-26-54 Tamatsukuri, Chuo-ku, Osaka 540-0004, Japan

Reviews

The editors invite reviews of books, tests, teaching systems, and other publications in the field of language education. A list of publications that have been sent to JALT for review is published monthly in *The Language Teacher*. Please send submissions, queries, or requests for books, materials, and review guidelines to:

Steve Cornwell, *JALT Journal* Interim Reviews Editor
Osaka Jogakuin Junior College
2-26-54 Tamatsukuri, Chuo-ku, Osaka 540-0004, Japan

Japanese-Language Manuscripts

JALT Journal welcomes Japanese-language manuscripts on second/foreign language teaching and learning as well as Japanese-language reviews of publications. Submissions must conform to the Editorial Policy and Guidelines given above. Authors must provide a detailed abstract in English, 500 to 750 words in length, for full-length manuscripts and a 100-word abstract of reviews. Refer to the Japanese-language Guidelines for details. Please send Japanese-language manuscripts to:

Sayoko Yamashita, *JALT Journal* Japanese Language Editor
Tokyo Medical and Dental University
International Student Center
2-3-21 Kanda-Surugadai, Chiyoda-ku
Tokyo 101-0062, Japan

Address for Inquiries about Subscriptions, Author Reprints, or Advertising

JALT Central Office
Urban Edge Building 5F
1-37-9 Taito, Taito-ku, Tokyo 110-0016, Japan
Tel.: 03-3837-1630; Fax: 03-3837-1631
(From overseas: Tel.: 81-3-3837-1630; Fax: 81-3-3837-1631)
website: <http://www.jalt.org>

日本語論文投稿要領

JALT Journalでは日本語で執筆された論文、研究報告、実践報告、書評等を募集しています。

文体：一般的な学術論文のスタイルを用い、章立ての仕方や参考文献のデータの書き方などは、*Publication Manual of the American Psychological Association* (5th ed.) の定める方式に合わせて下さい。不明の場合は、JALT Journal の日本語論文を参考にするか、日本語編集者までお問い合わせ下さい。なお、JALT Journal の読者は現場の教師が主なので、特殊な専門用語や統計的手法は、わかりやすく定義するか説明を加えるなどして下さい。

原稿：長さは、参考文献リストも含め18,000字（書評の場合は 1,500字）以内です。A4の用紙に横書きで、1行40字、1ページ30行で印刷して下さい。手書きの原稿は受け付けません。図表をいれる場合は、JALT Journal のページのサイズに合わせて（縮小コピー可）、本文とは別に作成して下さい。執筆者の名前や所属機関名は、原稿には書かないで下さい。原稿締切日は定めていません。いつでも受け付けます。

提出するもの：

- 原稿3部（執筆者名は書かないこと）
- 論文タイトル、執筆者の名前と所属機関名および連絡先（住所、電話番号、ファックス、e-mail アドレス）を書いた表紙
- 400字以内の和文要旨
- 英文のタイトルと、500～750語の英文要旨（書評の場合は100語程度の英文要旨）
- 100字以内の執筆者略歴
- MS-Word で保存したファイル（マッキントッシュ使用の場合はrtf ファイルで保存したもの）

査読：編集委員会が投稿要領に合っているかどうかを確認したあと、少なくとも二人の査読者が査読を行います。査読者には執筆者の名前は知らされません。査読の過程では特に、原稿が JALT Journal の目的に合っているか、言語教育にとって意味があるか、独創性はあるか、研究計画や方法論は適切か等が判定されます。査読は通常三か月以内に終了します。

注意：JALT Journal に投稿する原稿は、すでに出版されているものや他の学術雑誌に投稿中のものは受理できません。JALT Journal は、そこに掲載されるすべての論文に関して国際著作権協定による世界初出版権を持ちます。なお、お送りいただいた原稿は返却しませんので、控を保存して下さい。

投稿原稿送り先またはお問い合わせ：

〒101-0062 東京都千代田区神田駿河台2-3-21
 東京医科歯科大学留学生センター
 JALT Journal 日本語編集者 山下早代子
 電話・Fax: 03-5283-5861
 yama@tmd.ac.jp

JALT Journal 第26巻 第1号

2004年4月20日 印刷

2004年5月1日 発行

編集人 ドナ・タツキ

発行人 ジェイムス・スワン

発行所 全国語学教育学会事務局

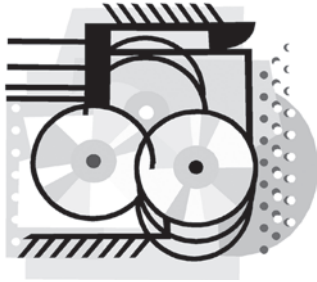
〒110-0016 東京都台東区台東1-37-9 アーバンエッジビル5F

TEL (03)3837-1630; FAX (03)3837-1631

印刷所 コーシンシャ株式会社

〒530-0043 大阪市北区天満1-18-4天満ファーストビル301 TEL (06) 6351-8795

Publications on CD ROM Available from JALT



The Language Teacher on CD-ROM *browse • search • learn*

The Language Teacher

- ▶ **Millennium TLT: Volumes 23 & 24 (1999-2000).** ¥1,500.
- ▶ **Episode 2: Volumes 11 through 18 (1986-1993).** Features more than 8000 pages of TLT and hundreds of articles. JALT Members: ¥3,500. Nonmembers: ¥4,000.
- ▶ **Episode 1: Volumes 1-10 (1976-1985).** JALT Members: ¥3,500. Nonmembers: ¥4,000.

Conference Proceedings

- ▶ **On JALT2000: Towards the New Millennium.**
Proceedings of the 26th annual JALT conference. JALT Members: ¥3,500. Nonmembers: ¥4,000. (Includes bonus: *PAC Journal*, Vol. 1, a journal for language teachers in Asia.)
- ▶ **On JALT99: Teacher Belief, Teacher Action.**
Proceedings of the 25th annual JALT conference. JALT Members: ¥3,500. Nonmembers: ¥4,000.
- ▶ **Conference Proceedings Archive CD, 1995-1998.**
Proceedings of the 21st to 24th annual JALT conferences. JALT Members: ¥3,500. Non-members: ¥4,000.

Other Publications

- ▶ **JALT Applied Materials: Second Language Acquisition Research in Japan.**
15 articles on the state of SLAR in Japan. ¥2,000.

Ordering

To Order: Use the postal cash transfer form at the back of this issue of *TLT*. Write the CD ROM title in the "Other" line. Credit card payment also accepted. Domestic and Overseas orders may be made by VISA or MasterCard. There is an additional ¥500 shipping and handling charge for overseas orders. Visit <www.jalt.org/main/shop> to download an order form, or please contact:

JALT Central Office, Urban Edge Bldg. 5F, 1-37-9 Taito, Taito-ku,
Tokyo 110-0016 JAPAN
TEL: 03-3837-1630; FAX -1637
Email: jalt@gol.com

Coming Soon on Archival CDs

Episode 3: JALT Applied Materials; JALT Journal

Release dates to be announced.

Windows and Macintosh compatible. Requires Adobe Acrobat Reader 4.0 or later, Acrobat e-Book Reader, or PDF compatible word processor.