A New World of English,

Through A Child’s Eyes

*Ready for English* is a lively and stimulating primary ESL learning system that helps the teacher approach English language learning from a child’s point of view.

This widely-acclaimed, field-tested program works because it utilizes the principles of natural, whole language learning—activities loaded with visual impact and physical involvement, plus a cast of real-life characters from diverse backgrounds in situations that students identify with and relate to.

*Ready for English* comes packaged in a sturdy, attractive storage cupboard, so everything’s in one place: teachers manuals, placement tests, colorful student manipulatives and visuals, song cassettes, student readers and the friendly, furry hand-puppet that encourages role-playing in fun situations.

*Ready for English* is ready for you. Call us or write us for our full-color brochure.

National Textbook Company
a division of *NTC Publishing Group*

4255 W. Touhy Avenue Lincolnwood, Illinois 60646-1975 U.S.A.
1-800-323-4900 (In Illinois 708-679-5500) FAX: 708-679-2494

AD0110
### ARTICLES

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>There is No Best Method—Why?</td>
<td>161</td>
<td>(10-25)</td>
</tr>
<tr>
<td>N. S. Prabhu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strategies of Unsuccessful Language Learners</td>
<td>177</td>
<td>(26-47)</td>
</tr>
<tr>
<td>Roberta J. Vann and Roberta G. Abraham</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Discourse of Computer-Assisted Language Learning:</td>
<td>199</td>
<td>(48-74)</td>
</tr>
<tr>
<td>Toward a Context for Descriptive Research</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carol Chapelle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicting Success for International Teaching Assistants in a U.S. University</td>
<td>227</td>
<td>(76-92)</td>
</tr>
<tr>
<td>George Yule and Paul Hoffman</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading-Writing Relationships in First and Second Language</td>
<td>245</td>
<td>(94-115)</td>
</tr>
<tr>
<td>Joan Eisterhold Carson, Patricia L. Carrell, Sandra Silberstein,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barbara Kroll, and Phyllis A. Kuehn</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Relationship Between Overall Reading Comprehension and</td>
<td>267</td>
<td>(116-141)</td>
</tr>
<tr>
<td>Comprehension of Coreferential Ties for Second Language</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Readers of English</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marjorie C. Demel</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### REVIEWS

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recent Publications on Statistics, Language Testing, and Quantitative Research Methods: II</td>
<td>293</td>
</tr>
<tr>
<td>Fundamental Considerations in Language Testing</td>
<td></td>
</tr>
<tr>
<td>Lyle Bachman</td>
<td></td>
</tr>
<tr>
<td>Assessing Achievement in the Learner-Centred Curriculum</td>
<td></td>
</tr>
<tr>
<td>Geoff Brindley</td>
<td></td>
</tr>
<tr>
<td>Testing for Language Teachers</td>
<td></td>
</tr>
<tr>
<td>Arthur Hughes</td>
<td></td>
</tr>
<tr>
<td>Reviewed by Liz Hamp-Lyons</td>
<td></td>
</tr>
<tr>
<td>Breaking Rules: Generating and Exploring Alternatives</td>
<td>301</td>
</tr>
<tr>
<td>John F. Fanselow</td>
<td></td>
</tr>
<tr>
<td>Reviewed by Ruth Spycher</td>
<td></td>
</tr>
</tbody>
</table>
BOOK NOTICES  305

Literature in the Language Classroom: A Resource Book of Ideas and Activities  Joanne Collie and Stephen Slater
Whole Language Strategies for ESL Students  Gail Heald-Taylor
Language, Learners and Computers  John Higgins
Computers in the Language Classroom  Robert M. Hertz
Longman English Grammar  Louis G. Alexander
The Origin of Writing  Roy Harris

BRIEF REPORTS AND SUMMARIES

Attitudes Toward “Cheating” Behaviors in the ESL Classroom  313
Phyllis Kuehn, Douglas J. Stanwyck, and C. L. Holland
Adult ESL Students’ Management of Dialogue Journal Conversation  317
Martha R. Dolly

THE FORUM

The Dirty Laundry of ESL Survey Research  323
Joy Reid

Information for Contributors  339
Editorial Policy
General Information for Authors

Publications Received  343
Publications Available from the TESOL Central Office  347
TESOL Membership Application  376
is an international professional organization for those concerned with the teaching of English as a second or foreign language and of standard English as a second dialect. Benefits of membership include four issues of the TESOL Quarterly and six issues of the TESOL Newsletter a year, plus a variety of other services and opportunities for professional development. Each year TESOL holds a Convention in the spring and a Summer Institute. TESOL also publishes and distributes a wide range of professional resource materials, which are available to members at reduced rates. A TESOL Membership Application is printed in every issue of the Quarterly. For further information, contact the TESOL Central Office, Suite 300, 1600 Cameron Street, Alexandria, Virginia 22314 U.S.A., (703) 836-0774.

TESOL QUARTERLY is published in Spring, Summer, Autumn, and Winter. Publisher’s representative is Helen Kornblum, Director of Publications. Subscriptions are not sold without membership. Communications regarding back issues, single copies, and permission to reprint should be addressed to the TESOL Central Office, Suite 300, 1600 Cameron Street, Alexandria, Virginia 22314 U.S.A., (703) 836-0774. Contributions should be sent to the Editor or the appropriate Section Editors at the addresses listed in the Information for Contributors, which is included in each issue of the Quarterly.

TESOL NEWSLETTER is published in February, April, June, August, October, and December. The Newsletter contains organizational news and announcements, Affiliate and Interest Section news and information, book reviews, conference information and reports, short refereed articles on current classroom practices, and general information. The Newsletter is available only through membership in TESOL or by subscription through a regional affiliate of TESOL. All contributions should be directed to Dr. William Powell, Editor, TESOL Newsletter, University of Southern Mississippi, Southern Station Box 5038, Hattiesburg, Mississippi 39406-5038 U.S.A., (601) 266-4964.

ANNOUNCEMENTS should be sent directly to the Editor, TESOL Newsletter, two months prior to the month of publication and must be received by the first of that month (e.g., February 1 for the April issue). The TESOL Quarterly does not publish announcements.

ADVERTISING in all TESOL publications is arranged by Helen Kornblum, TESOL Central Office, Suite 300, 1600 Cameron Street, Alexandria, Virginia 22314 U.S.A., (703) 836-0774.

OFFICERS AND EXECUTIVE BOARD 1990-91

President
ROBERT B. KAPLAN
University of Southern California
Los Angeles, CA

First Vice President
LYDIA STACK
Newcomer High School
San Francisco, CA

Second Vice President
SUZANNE M. GRIFFIN
Office of the State Superintendent of Public Instruction
Olympia, WA

Executive Director
RICHARD OREM
Alexandria, VA

Richard Allwright
University of Lancaster
Lancaster, U.K.

James D. Brown
University of Hawaii
Honolulu, HI

D. Scott Enright
Georgia State University
Atlanta, GA

Janet C. Fisher
California State University
Los Angeles, CA

Sergio Gaitan
Teachers College of Columbia University
New York, NY

Fred Genesee
McGill University
Montreal, Quebec, Canada

Elke V. Hamayan
Illinois Resource Center
Des Plaines, IL

Marjorie P. Knowles
Sawyer College
San Jose, CA

Carol Kreidler
Georgetown University
Washington, DC

Jean McConochie
Pace University
New York, NY

G. T. (Tippy) Schwabe
University of California
Davis, CA

Linda Tobash
LaGuardia Community College
City University of New York
Long Island City, NY

Rita Wong
San Francisco State University
San Francisco, CA
Editor’s Note

With this issue we inaugurate an editorial policy that modestly acknowledges the variety of written Englishes. Readers will note that we have elected not to “Americanize” British spelling or usages that would be recognizable to readers in North America. This policy raises unresolved issues concerning style and comprehension. We look forward to these challenges, believing that a diversity of voices will repay the effort.

Also with this issue, to provide a more comprehensive and informative Table of Contents, we introduce individual listings of volumes announced in the Book Notices section.

Please note that Brief Reports and Summaries Editor Gail Weinstein-Shr will be on leave from Temple University during the academic year 1990-1991. She will be a Visiting Scholar at the University of Massachusetts at Amherst. Effective immediately, all manuscripts for this section should be sent to the address listed in the Information for Contributors section.

In this Issue

Each of the articles in this issue of the TESOL Quarterly seeks greater precision in the way we conceptualize and pursue our work. The first article attempts to bring clarity to the often unexplored statement, “There is no best method.” The second uses a case-study approach to test critical assumptions about unsuccessful language learners. The third argues a need for comparability in research on computer-assisted language learning. Other articles in this issue address the selection of international students for teaching responsibilities, issues in literacy development, and second language reading comprehension. Each of the contributions to this issue takes an important turn in an on-going conversation about research and teaching.
N. S. Prabhu examines the substance of the statement “There is no best method.” He interrogates several possible explanations of the statement, noting that the most helpful considers misguided the very notion of good and bad methods. Prabhu speculates that “there is a factor more basic than the choice between methods, namely teachers’ subjective understanding of the teaching they do,” which he terms teachers’ sense of plausibility. A specialist who advocates a method can be said “to have articulated a fuller and more communicable articulation of a particular sense of plausibility.” A method, then, is simply “a highly developed and highly articulated sense of plausibility . . . not good or bad in any objective sense.” When teachers’ sense of plausibility is engaged, they can create classrooms that are active and alive.

Roberta Vann and Roberta Abraham explore strategies of two unsuccessful language learners as they completed four activities (an interview, a verb exercise, a cloze passage, and a composition). Task demands were delineated and learner strategies were disclosed by analyzing both think-aloud protocols and task products. Contrary to previous claims that unsuccessful learners use fewer strategies, these learners were active strategy users, though they sometimes applied strategies inappropriately. The authors note that “this research provides evidence for the importance of case studies in verifying critical assumptions about second language learning.”

Carol Chapelle notes that current research on computer-assisted language learning (CALL) “suffers from the same limitations as early research on classroom instruction”: There is a lack of precision and systematicity in describing CALL activities that precludes comparability across studies and between classroom and computer research. She argues that the discourse system of classroom interaction developed by Coulthard and Sinclair (1975) “provides the necessary elements and structures to describe CALL discourse, analyze data from student-computer interaction, and compare CALL activities with other (classroom) activities.”

George Yule and Paul Hoffman use TOEFL and GRE scores submitted at the time of application to predict which international graduate students will eventually be recommended to assume teaching duties on the basis of their performance in an ITA preparation course. Students who were not granted teaching assistantships were found, on average, to have significantly lower TOEFL and GRE Verbal scores than those who were recommended for teaching positions. These results suggest that TOEFL and GRE scores can act as predictors of students’ eventual success in presenting instructional material in English. However, Yule and Hoffman also note the responsibilities entailed by results that indicate that a substantial proportion of international graduate students will require English instruction after entering the university: “If U.S. universities intend to continue recruiting the brightest and best of the
world’s international graduate students to support university teaching and research missions, they should plan to provide more extended periods of adjustment and ESL training for a substantial proportion of those students."

• Joan Eisterhold Carson, Patricia Carrell, Sandra Silberstein, Barbara Kroll, and Phyllis Kuehn investigate reading-writing relationships in first and second language for Chinese and Japanese speakers. "The results indicate that literacy skills can transfer across languages, but the pattern of this transfer differs for the two language groups. It also appears that reading ability transfers more easily from L1 to L2 than does writing ability, and that the relationship of reading and writing skills varies for the two language groups." The study suggests that "L2 literacy development is a complex phenomenon for already literate adult second language learners." Different patterns of L2 literacy acquisition may involve variables such as L2 language proficiency, L1 and L2 educational experience, and differences in literacy and cultural practices.

• Marjorie Demel investigates the relationship between overall reading comprehension and the comprehension of coreferential pronouns, which are anaphoric in nature, i.e., "interpretable only through relationships that are presupposed in the text." The results suggest that misunderstanding coreferential ties results from difficulty in interpreting the elements of text to which these pronouns refer. This difficulty seems often to arise from inadequate knowledge of the target culture. Demel argues that second language students "need help both in locating pronominal referents and in developing strategies for interpreting these."

Also in this issue:

• Reviews: In the second part of a survey review of recent publications on statistics, language testing, and quantitative research methods, Liz Hamp-Lyons reviews Lyle Bachman’s Fundamental Considerations in Language Testing, Geoff Brindley’s Assessing Achievement in the Learner-Centred Curriculum, and Arthur Hughes’ Testing for Language Teachers; Ruth Spycher reviews John F. Fanselow’s Breaking Rules: Generating and Exploring Alternatives.

• Book Notices

• Brief Reports and Summaries: Phyllis Kuehn, Douglas Stanwyck, and C. L. Holland document attitudinal differences toward “cheating” behaviors among Spanish-speaking, Arabic-speaking, and English-speaking students, noting that “what is cheating in one culture may have an entirely different value in another”; Martha Dolly investigates native-speaker/nonnative-speaker interaction in the conversation management of dialogue journal writing.
The Forum: Joy Reid argues that “most research has some ‘dirty laundry’ that has been eliminated and/or concealed before publication.” In an effort to provide novice researchers a sense of the research process, and to assist others working to design survey instruments for an ESL population, Reid presents a retrospective report of developing and norming a survey instrument for ESL students.

Sandra Silberstein
There Is No Best Method—Why?

N. S. PRABHU
National University of Singapore

This paper examines the possible substance of a statement, often heard but rarely explained, in the discussion of language teaching, namely, that there is no best method. Three possible explanations of the statement are discussed, based on a broad interpretation of the term method: (a) that different methods are best for different teaching contexts; (b) that all methods are partially true or valid; and (c) that the notion of good and bad methods is itself misguided. The first two explanations are shown to be unhelpful; the third is discussed in some detail, with an exploration of the concept, teachers’ sense of plausibility.

It is uncommon these days to have a sustained discussion on language teaching without someone at some point declaring that there is no best method, or words to that effect. Such a declaration usually occurs at a late stage in an indecisive debate about different methods, and has the general effect of altering the orientation of the debate itself, somewhat abruptly. It also carries the ring of an incontrovertible statement—or a statement so tolerant and reconciliatory in spirit that to dispute it would be professionally churlish. As a result, one rarely sees a detailed examination of what it might mean to say that there is no best method.

I think it helpful to see the statement as having an illocutionary as well as a propositional meaning. As an illocutionary act, it seeks to terminate a debate without reaching any substantive conclusion, and it does so not by admitting defeat in the effort to reach a conclusion, but by appearing to raise the debate itself to a higher level, thus helping to save professional face all round. It suggests not just that proponents of different methods agree to disagree, but that they give up their pursuit of agreement and disagreement as being unproductive. It manages to reconcile conflicting views by defining a position to which all can honorably subscribe, and by striking a philosophical note on which enlightened discussion can properly end. It is thus a convenient device for legitimizing nonresolution of methodological issues.
The propositional content of the statement, on the other hand, calls for a great deal of clarification and substantiation, in terms of the sense in which a method can be considered good, the reason why no method can be thought or shown to be better than others, the consequence of that position for the theory and practice of pedagogy, the revaluation it demands of the assumptions currently underlying conflict and debate between methods, and the forms in which debate or discussion may still be of use from the new perspective. What is involved is not just the termination of a discussion, but the beginning of a new one—or rather the beginning of a new phase in the discussion, as a positive outcome of the earlier phase. It is this propositional content of the statement that is the concern of this paper; I hope to clarify the sense in which there can be said to be no best method of language teaching.

First, however, a word about the term method. I use the term inclusively, to refer both to a set of activities to be carried out in the classroom and to the theory, belief, or plausible concept that informs those activities. The reader will thus find one or both of these aspects the focus of particular parts of the discussion. I consider this “global” interpretation of the term appropriate to analysing a statement that is equally global in spirit.

There are, I think, three general lines of argument that can be advanced in support of the statement to be discussed. These are examined in turn.

**IT ALL DEPENDS ON THE TEACHING CONTEXT**

If those who declare that there is no best method are asked why, the most immediate and frequent answer is likely to be “Because it all depends,” meaning that what is best depends on whom the method is for, in what circumstances, for what purpose, and so on. That there is no best method therefore means that no single method is best for everyone, as there are important variations in the teaching context that influence what is best. The variations are of several kinds, relating to social situation (language policy, language environment, linguistic and cultural attitudes, economic and ideological factors, etc.), educational organisation (instructional objectives, constraints of time and resources, administrative efficiency, class-size, classroom ethos, etc.), teacher-related factors (status, training, belief, autonomy, skill, etc.), and learner-related factors (age, aspirations, previous learning experience, attitudes to learning, etc.). There have been several attempts to categorize such variables systematically and comprehensively (e.g., see Brumfit, 1984), but even the brief and random listing above shows that they
are at different levels of generality, as well as of discreteness and tangibility. Moreover, recent and current work in the field seems to be adding new factors and categories to the inventory, in the form of varied learning styles, communication strategies, personality factors, and psychological processes. Together with slightly earlier work on variations in learner’s needs and purposes, this seems to suggest that variability on such dimensions is infinite, thus strongly challenging the notion that any given method can be good for everyone. “It all depends,” and what it depends on is a vast number of things.

Notice, however, that to say that no single method is best for everyone is also to say that different methods are best for different people—or for different teaching contexts. This implies that, for any single teaching context, there is in fact a method that is best and, further, we are able to determine what it is. If we are unclear or in disagreement about what method is best for a specific context, there is need for discussion, debate, and interaction between differing perceptions; we should be seeking to further pedagogic debate rather than to terminate it with a face-saving formula. It is perhaps not without significance that statements like “There is no best method” are made most often as defensive postures—as ways of saying not only that one does not agree with what is being argued, but that one refuses to engage any further in the argument. Such statements succeed in preserving the conversational peace, but cause a loss of the productive potential of professional debate. For such statements to act as a contribution to debate, we will need to interpret them as proposing only a narrowing of the scope to some single teaching context, so that dependencies between contextual variables and methodological options might be explored, and a search for the best method might be continued with greater focus. The statement that there is no best method would then be an assertion, not of the futility of looking for the best method, but of the desirability of asking what method is best for some specific context; contextual variability would serve not as a means of avoiding methodological issues, but as a possible new approach to resolving them.

I think it is important to realise how complex it can be to determine dependencies between contextual factors and instructional methods. To begin with, many of the factors that are discussed are neither easy to identify nor simple to assess: They tend suspiciously to fit different slots in different taxonomies. Formal environment, for instance, may refer to classroom learning, as against learning through social exposure, or to the formal school system, as against private language instruction, or to relative
formality in teacher-learner relations, as against informal relations, or even to teacher-fronted activities, as against group work among learners. *Motivation* may mean anything from future career ambitions, to a desire for group approval, to a passing interest in a particular classroom activity. What are identified as attitudes are seldom clear-cut or unmixed or even stable. When we come to factors like preferred learning styles, sociocultural influences, or personality factors, we are faced with unclear and overlapping distinctions, and are therefore forced to simplify and stereotype, often in preconceived ways. Further, even when some contextual factors are clearly identifiable, their consequence for instructional procedures can be far from clear. Do older learners need a different method of teaching from younger ones, and if so, how fundamentally different? Is a change in method more likely to succeed in the hands of experienced teachers, or less? If there is a mismatch between official language policy and learners' personal goals, which should have what weight in the choice of an instructional method? If we identify certain learning strategies that learners naturally tend to employ, do we conclude, for that reason, that they are good learning strategies? If not, are learners likely to learn more by following their own strategies, which may be less than good in our view, or by adopting strategies we consider more conducive to learning, though they may go against such natural tendencies? If earlier experience has conditioned learners or teachers to certain perceptions of learning and teaching, does that constitute an argument against change, or indicate a greater need for change?

The point I am making is not just that our knowledge is uncertain at this time; the more important point is that it is only when we can show a relationship between a contextual factor and a methodological decision that the contextual factor becomes significant for pedagogy. What we need is not just an identification and projection of variation but, equally, some way of determining which form of variation matters to instruction and how, and which does not. If we look for variation merely on the assumption that the teaching context matters for teaching methodology, we are sure to find indefinite variation on many dimensions, thus making it impossible to justify any instructional method for any single group of learners. If all physiological variation among individuals (including fingerprints) were assumed to call for matching differentiation in medical treatment, no medical practice would be justifiable.

It is, of course, possible to obviate the problem of relating contextual factors to instructional methods by giving contextual factors a central role in pedagogy and treating instructional methods as a kind of logical derivation from them. This is the move
often advocated: from a preoccupation with teaching methods to an
effort at “curriculum development” or “course design.” The assump-
tion is that, with such a move, decisions concerning methods will
either be rendered unnecessary or play only a small part in the enter-
prise. As one applied linguist puts it, “The important issues are not
which method to adopt but how to develop procedures and instruc-
tional activities that will enable program objectives to be attained”
(Richards, 1985, p. 42). The procedures envisaged are those of situ-
tional analysis, needs analysis, analysis of “authentic” samples of
target language use, surveys of opinions and attitudes, estimates of
resources, etc.—that is to say, compilations of different kinds of in-
formation about learners, teachers, the school, and the society, with a
view to determining instructional objectives. Instructional proce-
dures are seen to follow, or be easily determinable, from the profile
of contextual factors and the statement of objectives. This is, in
effect, a kind of discovery procedure for methods: That method is
best, it seems to say, which results from a careful implementation of
the procedure, the soundness of the method being guaranteed by the
soundness of the procedure leading to it.

There is, however, a price to pay for this simplification of
pedagogy. The instructional procedures most directly derivable
from a specification of needs, wants, and objectives are those of
supplying to learners the relevant tokens of language, or getting
them to rehearse target language behaviour in simulated target
situations. Any concept of developing in learners a more basic
capacity for generating tokens of language when needed, or for
adapting to unforeseen target language behaviour as necessary,
leads one toward ideas about the nature of language ability and the
process of language acquisition—complex methodological issues
that the discovery procedure seeks to avoid. Besides, a more
elaborate analysis of contextual factors results in a correspondingly
larger set of criteria to be met by instructional content or procedure,
and the larger the set of criteria to be met, the fewer the choices
available in meeting them. Language instruction that attempts to
cater directly to social objectives, learning needs, target needs,
learners’ wants, teachers’ preferences, learning styles, teaching
constraints, and attitudes all round can end up as a mere assemblage
of hard-found pieces of content and procedure—a formula that
manages, with difficulty, to satisfy multiple criteria and therefore
cannot afford to let itself be tampered with. What sets out to be a
widening of attention to varied educational considerations can thus
end up as an abridgement of choice and flexibility in the practice of
pedagogy. In avoiding adherence to a single method, one will have
arrived instead at a single, fixed, teaching package.
Avoiding adherence to a single method has a certain ideological aura to it. It suggests liberation from a monolithic mould, a refusal to be doctrinaire, an espousal of plurality. It is, however, also a denial of the role of understanding in language pedagogy, which is necessarily a matter of ideation: We understand something when we have a set of ideas or principles that cohere to make up a conceptual model, or theory. Although theory is only an abstraction, it is about the only instrument we have (at any rate, the most accessible instrument) for making sense of complex phenomena and conveying that sense to one another. Theory, as we know, arises not from a cataloging of diversity, but from a perception of unity in diverse phenomena—a single principle, or a single system of principles, in terms of which diversity can be maximally accounted for. If the theories of language teaching (that is to say, methods) that we have at present fail to account sufficiently for the diversity in teaching contexts, we ought to try to develop a more general or comprehensive (and probably more abstract) theory to account for more of the diversity, not reject the notion of a single system of ideas and seek to be guided instead by diversity itself. Pointing to a bewildering variety of contextual factors as a means of denying the possibility of a single theory can only be a contribution to bewilderment, not to understanding.

There is also the fact that a concentration on dissimilarities between teaching contexts is likely to obscure similarities between them; in the case of language learning, we are dealing with a human ability that constitutes a defining characteristic of the species. It is true that variations in social situation, institutional organisation, individuals’ histories, attitudes, or intentions can all have the effect of limiting or extending opportunities for desired forms of pedagogic action. However, to imply that they call for a matching differentiation in pedagogic theories is to make the very large claim that the process of language acquisition—a basic human attribute—itself varies according to contextual factors.

THERE IS SOME TRUTH TO EVERY METHOD

I have discussed in some detail one form of substantiation of the statement that there is no best method, namely, that different methods are best for different teaching contexts. A different form of substantiation is also heard fairly frequently, namely, that there is some truth (or value or validity) to every method—or, at any rate, to several different methods—even though the methods may be conceptually incompatible. This, as we know, is an argument for eclecticism in language pedagogy—not an argument that different
contexts should use different methods, but an argument that the same context should use a number of different methods, or perhaps parts of different methods. There is an immediate appeal to common sense in this stance: If every method is a partial truth, then it seems clear that none represents the whole truth; to adopt any single method is to settle for much less than one can get by adopting all or several of them.

As a comment on our state of knowledge at this time (or indeed at any time), the suggestion that no method contains more than a partial truth is clearly unexceptionable. We continue to engage in the professional activity of research, concept development, discussion, and debate because all of our understanding of language learning and teaching is at best partial, and for any of us operating with a theory, which can represent only a partial truth, it remains entirely possible that other theories represent partial truths as well. However, this philosophical perception of imperfect knowledge does not help us to see which theory represents which part of the truth—or which part of a given theory represents the truth. Were we able to see that, we would no longer be operating with a theory of which only a part is known to represent the truth, and missing those parts of other theories that are known to represent other parts of the truth. Our knowledge would make a leap toward the whole truth—whatever that might mean. The fact, however, is that the understanding one has of a phenomenon at a given time—the theory one is operating with—represents for one the whole truth. Other theories are true to the extent they share the understanding represented by one’s own theory. In this sense, each theory can claim to represent the partial truths of other theories: If there is an overlap of understanding between it and certain parts of other theories, then it can be said to contain those parts of other theories that constitute partial truths.

But the statement that there is some truth to every method needs to be seen not just as an epistemological observation, but as a plea for an eclectic blending of all or several methods. Now, any such blending of different methods is either done with a perception of what is true about each method, or it is done without any such discrimination. If there is a perception of which part of what method is a partial truth, then that perception constitutes a theory, which happens to have an overlap of understanding with various other theories. It therefore represents a method, which is like any other method, with an overlap of understanding with others. There is no reason to think, on the strength of its being a blend, that it has any more of the truth than any other method. It is simply one of the methods that share some of their concepts or procedures with other
methods; how much of the truth it represents is a matter not of how much blending it does, but of what particular perception makes the blending possible. What we have, therefore, is not an eclectic blend, but a different method—or else, all methods which happen to have partial overlaps with others are equally eclectic.

If, on the other hand, an eclectic blending of different methods is done not with any particular perception of what parts of those methods represent the truth, but rather in the hope that whatever is true about them will be captured in the blending, then the eclectic blend does not constitute a method, but instead an act of gambling or a hedging of bets: It can only have treated all parts of different methods as being equally likely to be true or untrue, and, as a result, there is as strong a possibility of its being a blend of the untruthful parts of different methods as there is of its representing the truthful parts (whatever notion of truth and untruth we may care to employ). Further, such indiscriminate blending of methods adds nothing to our pedagogic understanding, since it offers no perception of what may be true about which method. It simply plays it safe—as safe from truth as from untruth. An eclectic blending that constitutes a form of pedagogic understanding at least offers us an additional method, though it makes an unjustified claim to being more than an additional method; an eclectic blending that does not constitute an additional method in that sense leads us away from any furtherance of understanding, while offering us a chance at what may be called “truth by accident.”

**WE NEED TO RETHINK WHAT “BEST” MIGHT MEAN**

Let me now turn to a third possible way of substantiating the statement that there is no best method. This is that we have no adequate notion of what “best” might mean—or that the notion of good and bad needs to be reexamined and clarified.

A prevalent notion of the best method is that it is the method that yields the best results in terms of learning outcomes. Since the aim of all teaching is to bring about as much learning as possible as quickly as possible, it seems self-evident that teaching methods should be judged by the amounts of learning they can lead to, in a given period of time. This appears to call for a comparison of methods and a quantification of learning outcomes, through well-designed, controlled experiments, in keeping with the spirit of objective, scientific enquiry. It is true that such objective evaluation is so difficult to implement that all attempts at it in the past have resulted in a wider agreement on the difficulties of doing an evaluation than on the resulting judgement on methods. It is also
true that arguments have been put forward for a possible alternative to the experimental design. Nevertheless, we generally continue to assume, more or less consciously, that there is a method that is objectively the best, that it is in principle possible to demonstrate that fact, and that once demonstrated, the superiority of the best method will lead to its widespread acceptance in the profession. That is to say, we generally see ourselves as working to that ideal, on the tacit premise that what is unrealised is not necessarily unrealizable, and that all our professional endeavour is a form of progress toward it. Alternatives such as trying to construct comprehensive descriptions of methods—as "illuminative" evaluation (Parlett & Hamilton, 1977)—involve an abandonment of that ideal, thus threatening to disorient our professional thought. We prefer to retain the ideal as the basis of all our professional effort.

Seen in this context, the statement that there is no best method is a questioning of the current concept of the best method—an argument that the ideal of objective and conclusive demonstration is not only an unrealised one, but an inherently unrealizable one, and that working with such an ideal is unproductive for the pedagogic profession. Brumfit (1984), for instance, has strongly questioned the notion that teaching methods, which are essentially concerned with human interaction, can usefully be subjected to the processes of objective testing and prediction, which are part of the scientific method. He argues, in summary, (a) that a teaching method in operation is necessarily an embodiment of certain general pedagogic principles into a variety of specific contextual features (including participants’ psychological states); (b) that predictive testing of a method demands manipulation and control of the manifold contextual features; (c) that many of the contextual features are either difficult or impossible to control; and, most important, (d) that any success actually achieved in controlling contextual features will have only the effect of disembodied the method, as it were, of its actual, operational form, thus rendering the outcome of the testing inapplicable to the operation of the method in any specific context. Brumfit comments:

A claim that we can predict closely what will happen in a situation as complex as [the classroom] can only be based on either the view that human beings are more mechanical in their learning responses than any recent discussion would allow, or the notion that we can measure and predict the quantities and qualities of all these factors. Neither of these seems to be a sensible point of view to take. (pp. 18-19)

While Brumfit’s argument is based largely on the complexity of the pedagogic operation, it is also possible to point to the
complexity of assessing language attainment as such. Examining aspects of this latter complexity further underscores the futility of attempts to objectify method evaluation. To begin with, an important consideration for language teaching methods is the quality of learning to be promoted, as distinct from the quantity. The question of quality has been a recurrent concern for the profession through the ages, being conceptualised and verbalised variously as grammar in contrast to practice, knowledge in contrast to skill, explicit knowledge in contrast to implicit knowledge, accuracy in contrast to fluency, learning in contrast to acquisition, ability to display in contrast to ability to deploy, etc. There may be disagreements about how the different kinds of knowledge or ability are related to each other, but it is remarkable how, whenever a distinction is made between different forms of knowledge of a language, it is the less conscious, less observable, and less quantifiable form that is seen to be of greater value or validity. Objective evaluation of methods, however, necessarily relies on a quantification of learning outcomes, and therefore tends to measure the more quantifiable form of knowledge. This means that the more objective the evaluation is, the less likely it is to assess learning of the desired quality, and vice versa.

Second, a perception of language ability as an implicit form of knowledge is linked to a perception of its development as an internal, unobservable process that is organic rather than additive, and continuous rather than itemizable. This means that at any stage of the growth process, there is not only the growth achieved so far, but a potential for further growth achieved as a part of it—a potential that can be thought of in terms of inchoation or incubation. Our most ambitious effort at language testing can only hope to give us evidence on the actual growth achieved at the stage of testing, not on the potential generated for further growth, since knowledge in an inchoative state is even less accessible to elicitation and quantification than implicit knowledge as such. Again, not everyone in the profession may regard knowledge of a language as being equally implicit in nature or organic in its development, but the point is that an objective evaluation of methods is unlikely to be able to cope with concepts of implicitness and inchoation and, as a result, unlikely to provide widely acceptable decisions.

Third, an objective evaluation of methods is not just an assessment of learners’ language attainments; it also involves an objective attribution of the learning that has taken place to the teaching that has been done. However, the relationship between language teaching and language learning becomes less and less direct as one perceives language as being an implicit ability and an internal
development: What is less conscious and less observable is also less directly teachable. The more indirect the relationship is between teaching and learning, the more difficult it is to attribute any specific piece of learning to any specific piece of teaching. It is, ultimately, difficult to tell what learning has taken place as intended by the teaching, and what has taken place independent of it (or, indeed, what in spite of it). We have problems enough maintaining a subjective perception of general causation between teaching and learning in the development and discussion of particular methods. An evaluative comparison of different methods calls for a degree of objectivity and specificity in cause-effect relations that may well be unreasonable to expect in the field of language pedagogy.

More generally, the notion behind an objective evaluation of methods is that there is something in a method that is by itself— independent of anyone’s subjective perception of it—superior or inferior to what there is in another method. If some method were shown by such evaluation to be superior to all others, then that method would be expected to benefit all (or a large number of) classrooms, regardless of how it is subjectively perceived by the different teachers involved. A method, in this view, is a set of procedures that carries a prediction of results; the fulfillment of the prediction depends only (or mainly) on an accurate replication of the procedures, not on any perceptions of those who do the replication—rather in the way the replication of a procedure in chemistry yields the predicted result, regardless of the chemist’s thoughts or feelings about it. No doubt the idea looks fairly absurd when put in this form: It reduces teaching to a faithful following of highly specified routine—something of a pedagogic ritual. I am, however, unable to see how a serious pursuit of objective method evaluation can be sustained without some such idea. The only alternative to it is to maintain that the method that is shown to be objectively superior will somehow carry with it the subjective perception that lay behind its development and, equally, that the perception concerned will then replace the differing perceptions of all the teachers who may be led to adopt that method on the strength of the objective evaluation. This implies, among other things, that teachers’ pedagogic perceptions are as easily replaceable as classroom procedures, an idea that could hardly be less absurd.

It is useful to ask why it looks absurd to suggest that a good teaching method can be carried out, without loss, as merely a routine. We find it necessary to think of good teaching as an activity in which there is a sense of involvement by the teacher. When we encounter an instance of really bad teaching, it is most often not a
case of the teacher following a method with which we disagree, but rather of the teacher merely going through the motions of teaching, with no sense of involvement. When a method considered to be good has been implemented on a large scale and later thought not to have “worked,” an important part of the reason identified has been that teachers followed the method “mechanically,” with no sense of understanding or identification. Indeed, the more “efficiently” a method is implemented (that is to say, with all possible measures to ensure that teachers will carry out the procedures envisaged), the more likely it is that mechanical teaching will turn out to be the main impediment to success.

Perhaps, then, there is a factor more basic than the choice between methods, namely, teachers’ subjective understanding of the teaching they do. Teachers need to operate with some personal conceptualisation of how their teaching leads to desired learning—with a notion of causation that has a measure of credibility for them. The conceptualisation may arise from a number of different sources, including a teacher’s experience in the past as a learner (with interpretations of how the teaching received at that time did or did not support one’s learning), a teacher’s earlier experience of teaching (with similar interpretations from the teaching end), exposure to one or more methods while training as a teacher (with some subjective evaluation of the methods concerned and perhaps a degree of identification with one or another of them), what a teacher knows or thinks of other teachers’ actions or opinions, and perhaps a teacher’s experience as a parent or caretaker. Different sources may influence different teachers to different extents, and what looks like the same experience or exposure may influence different teachers differently.

The resulting concept (or theory, or, in a more dormant state, pedagogic intuition) of how learning takes place and how teaching causes or supports it is what may be called a teacher’s sense of plausibility about teaching. This personal sense of plausibility may not only vary in its content from one teacher to another, but may be more or less firmly or fully formed, more or less consciously considered or articulated, between different teachers. It is when a teacher’s sense of plausibility is engaged in the teaching operation that the teacher can be said to be involved, and the teaching not to be mechanical. Further, when the sense of plausibility is engaged, the activity of teaching is productive: There is then a basis for the teacher to be satisfied or dissatisfied about the activity, and each instance of such satisfaction or dissatisfaction is itself a further influence on the sense of plausibility, confirming or disconfirming or revising it in some small measure, and generally contributing to
its growth or change. I also think that the greater the teacher’s involvement in teaching in this sense, the more likely it is that the sense of involvement will convey itself to learners, getting them involved as well and helping to create that elusive but highly regarded condition in the classroom: teacher-learner rapport. It is of course possible that other factors have a role, too, in the creation of rapport, such as learners’ own perceptions of learning and their interpretations of the teaching activity (Allwright, 1984). My point is that an engagement of the teacher’s sense of plausibility is a major condition for classroom rapport, whether or not it is the only condition. It is common to hear that learning is enhanced when learners enjoy classroom activity, but enjoyment is a broad notion and is often equated with some form of light entertainment interspersed with more serious activity. I think there is a form of enjoyment arising from teacher-learner rapport that is less conspicuous but more integral to classroom activity, and more truly productive of learning.

The picture of classroom activity that engages the teacher’s sense of plausibility is no doubt closer to an ideal than to a factual description of much of the teaching that actually goes on. But that does not detract from the suggestion I am making, namely, that that ideal is more worth our while to pursue than the notion of an objectively best method. The question to ask about a teacher’s sense of plausibility is not whether it implies a good or bad method but, more basically, whether it is active, alive, or operational enough to create a sense of involvement for both the teacher and the student. Mechanical teaching results from an overroutinisation of teaching activity, and teaching is subject to great pressures of routinisation. It is, after all, a recurrent pattern of procedures on regularly recurrent occasions. It is also a form of recurrent social encounter between teachers and learners, with self-images to protect, personalities to cope with, etc. And, like all recurrent social encounters, teaching requires a certain degree of routine to make it sustainable or even endurable. There are, in addition, varied feelings of adequacy or confidence among teachers, varied degrees of concern for maintaining status, stress of overwork, threat of peer comparisons or of expectations from superiors, etc., all of which can use the protection offered by role-defining routines. More fundamentally, there is the fact that learning by learners is essentially unpredictable and, consequently, it is an unacceptably high risk for teaching to have to justify itself in direct relation to learning: There is need for a way to claim that the teaching expected has been performed, though the learning expected may still not have occurred; and teaching defined as a routine precisely meets that need.
An active sense of plausibility is very difficult to maintain among such pressures on routinisation, and can easily become frozen, ossified, or inaccessibly submerged, leaving only a schedule of routines. When teachers profess to believe in some method they have been following—perhaps to the point of swearing by it—they may well be merely demonstrating how frozen their sense of plausibility is and, as a result, how insecure they feel against a threat to their teaching routines. When a teacher’s sense of plausibility is active and engaged in the teaching, it is necessarily open to change, however slowly or imperceptibly, in the process of the ongoing activity of teaching. Such teaching can perhaps be regarded as being “real,” in contrast to teaching that is mechanical. We can then say that a distinction between “real” and mechanical teaching is more significant for pedagogy than any distinction between good and bad methods. The enemy of good teaching is not a bad method, but overroutinisation.

If it is important for a teacher’s sense of plausibility to remain alive and therefore open to change—not frozen but fluid in some degree—then an important goal for the pedagogic profession is to seek ways in which the sense of plausibility in as many teachers as possible can be helped to remain as alive as possible, though necessarily in varied forms. It is true that the ongoing activity of teaching is itself a source of continual influence on a teacher’s sense of plausibility, thus helping to keep it alive, but we have noted how the ongoing activity of teaching is, at the same time, subject to varied pressures of routinisation, which can have a deadening effect on the sense of plausibility. A second source of influence on the sense of plausibility—perhaps the most important one outside the classroom—is interaction between different senses of plausibility. This interaction can arise from an articulation and discussion among teachers of one another’s pedagogic perceptions, from professional reading or writing, and in other, more or less formal, ways.

A specialist-level debate between different methods is, in fact, an interaction between different senses of plausibility, seeking to exert an influence on all those who participate in it (more or less overtly) through a process of sharing, sharpening, strengthening, weakening, changing, or helping to develop further the different forms of understanding involved. A specialist who advocates a method is on the same footing as a teacher who operates with a sense of plausibility, except that the specialist can be said to have achieved fuller and more communicable articulation of a particular sense of plausibility, perhaps in the course of wider (or longer or more intensive) interaction with other similarly well-articulated perceptions. If ossification is less likely to occur at the specialist’s level, it is only to
the extent that the specialist has more of a commitment to professional interaction and is more continually engaged in exploring and articulating some sense of plausibility. The resulting well-developed and well-articulated senses of plausibility (that is to say, methods) have value not as desirable replacements for many teachers’ senses of plausibility but for what may be called their power to influence—to invoke, activate, interact with, alter in some way, and generally keep alive—different teachers’ differing senses of plausibility, thus helping to promote and enlarge the occurrence of “real” teaching. A method, from this point of view, is not good or bad in any objective sense, but has more or less pedagogic power to influence teachers’ subjective understanding of teaching; and debates between different methods are important for the profession because they help to give expression and opportunity to the pedagogic power of different methods.

CONCLUSION

To summarise, if we regard our professional effort as a search for the best method which, when found, will replace all other methods, we may not only be working toward an unrealizable goal but, in the process, be misconstruing the nature of teaching as a set of procedures that can by themselves carry a guarantee of learning outcomes. To say that the best method, in this sense, varies from one teaching context to another does not help because it still leaves us with a search for the best method for any specific teaching context. To say that there is some truth to every method does not help either, because it still does not tell us which part of which method is true. Objective method evaluation has either to assume that methods have value for learning independent of teachers’ and students’ subjective understanding of them, thus perpetuating an unrealizable goal and reinforcing the misconstruction of pedagogy, or to try to take into account teachers’ subjective understanding of teaching, thus ceasing to be objectively evaluative. If, on the other hand, we view teaching as an activity whose value depends centrally on whether it is informed or uninformed by the teacher’s subjective sense of plausibility—on the degree to which it is “real” or mechanical—it becomes a worthwhile goal for our professional effort to help activate and develop teachers’ varied senses of plausibility. A method is seen simply as a highly developed and highly articulated sense of plausibility, with a certain power to influence other specialists’ or teachers’ perceptions. Perhaps the best method varies from one teacher to another, but only in the sense that it is best for each teacher to operate with his or her own sense of
plausibility at any given time. There may be some truth to each method, but only in so far as each method may operate as one or another teacher’s sense of plausibility, promoting the most learning that can be promoted by that teacher. The search for an inherently best method should perhaps give way to a search for ways in which teachers’ and specialists’ pedagogic perceptions can most widely interact with one another, so that teaching can become most widely and maximally real.

ACKNOWLEDGMENT

A version of this paper was presented as a plenary address to the 23rd annual meeting of the International Association of Teachers of English as a Foreign Language (IATEFL) held in Coventry, England, April 1989.

THE AUTHOR

N. S. Prabhu teaches applied linguistics at the National University of Singapore. His earlier work in India included the development of task-based language teaching and the concept of a procedural syllabus, in the context of what is generally known as the Bangalore Project.

REFERENCES


Strategies of Unsuccessful Language Learners

ROBERTA J. VANN and ROBERTA G. ABRAHAM
Iowa State University

Recent research on learning strategies has yielded conflicting findings and generated limited success in learner training. These problems may be rooted in inadequate knowledge of the actual strategies used by unsuccessful learners in contrast to what they report doing. The present study combines methods to probe the strategies of two unsuccessful learners—both Saudi Arabian women enrolled in an academically oriented intensive English program (IEP)—as they completed four activities (an interview, a verb exercise, a cloze passage, and a composition). After task requirements were determined, learner strategies were ascertained by analyzing think-aloud protocols and task products. These combined analyses offer a detailed and insightful picture of learner strategies, providing counterevidence for the claim that unsuccessful learners are inactive. When viewed through the task-demand model proposed here, these unsuccessful learners emerged as active strategy users, though they sometimes applied strategies inappropriately. The model also revealed fundamental differences in the approaches to problem solving used by learners who appear similar on the basis of simple strategy counts. This research provides evidence of the importance of case studies in verifying critical assumptions about second language learning.

Interest in language learning strategies (behaviors that learners engage in to learn a second/foreign language) and in attempts at remediating the strategies of unsuccessful language learners have blossomed in recent years (see, for example, Wenden & Rubin, 1987; O’Malley & Chamot, 1990; Oxford, 1990). Yet learner strategy research, for all its promise is, as Skehan (1989) points out, still “embryonic” (p. 98), with conflicting methods and results and few unequivocal findings. A few early attempts to train unsuccessful learners to use the strategies of their more successful peers resulted in improved performance (Cohen & Aphek, 1980; Hosenfeld, 1984). However, only slight gains were observed in a more recent and elaborate training study (O’Malley, Chamot, Stewner-Manzanares,
Further, the reported successes have all involved training for a specific task; generalization and transfer of trained strategies, found to be difficult to induce in first language settings (see, for example, Brown, 1980; Pressley, Levin, & Bryant, 1983; Scardamalia & Paris, 1985) have not been explored in second language research. Thus, even those most optimistic about the promise of learner training (Wenden, 1987a; O’Malley & Chamot, 1990) acknowledge the need for further study of its feasibility and effectiveness in second language learning.

One of the problems in developing effective strategy training may be the dearth of evidence concerning the actual strategies used by unsuccessful learners. The basis for training has typically been descriptions of the strategies of successful language learners (see, for example, Rubin, 1975; Stern, 1975; Naiman, Frohlich, Stern, & Todesco, 1978; and Chamot, 1987) on the assumption that poor learners lack these strategies. Only in a few studies (for example, Hosenfeld, 1976, 1984; Abraham & Vann, 1987; Chamot & Kupper, 1989) have unsuccessful language learners been observed.

Further, these descriptions of the “good” language learner have in most cases been based on teacher/researcher observations and learners’ generalized (often retrospective) self-reports. In some recent studies (see, for example, Chamot, 1987), interviews with learners have generated statements such as “I look up words I don’t know” or “I talk to native speakers as often as possible.” In others, (for example, Politzer & McGroarty, 1985), questionnaires on strategies and learning behaviors have been used. Little subsequent research has attempted to confirm how strategies are utilized as learners engage in actual tasks. Data collection methods that rely primarily on learners’ reported strategy use on questionnaires may account for certain puzzling findings. For example, O’Malley et al. (1985) noted that subjects in their study reported using more strategies for activities such as vocabulary learning and pronunciation than for complex activities involving analysis and inferencing. One would expect, however, that as learning tasks increase in difficulty, requiring more time and greater control of processing, more strategies rather than fewer would be needed. That strategies are reported more frequently for simple tasks suggests that the method of data collection itself may have skewed results.

As noted by Chamot (1987), this incomplete picture can be augmented by supplementing earlier research on the reported strategies of learners with detailed descriptions of how learners use strategies in specific situations. One means of obtaining such descriptions is by asking learners to “think aloud,” that is, to report what they are thinking as they are performing a task. Recently, cognitive psychologists such as Ericsson and Simon (1987) have...
argued that such methods of concurrent introspection are indispens-
able in probing many human thought processes that would otherwise
be inaccessible. Although this procedure, like introspective tech-
niques in general, is not without critics (see, for example, Seliger,
1983), forms of concurrent verbal introspection such as thinking
aloud are now recognized as essential tools in investigating language
processing (Cohen & Hosenfeld, 1981; Faerch & Kasper, 1987). When
combined with other methods, they may provide us with convergent
validity for certain findings, or force us to reexamine assumptions
about the role of strategies in language learning.

THE STUDY

The research reported here focuses on two unsuccessful language
learners to discover reasons for their lack of success in completing
an academic program. To obtain as unbiased and complete a
picture as possible, we followed the approach of methodological
triangulation (Grotjahn, 1987), combining the introspective think-
aloud technique with analysis of learners’ products.

This research is part of a larger study in which we set out to deter-
mine how successful and unsuccessful learners (as measured by the
relative speed with which they moved through an intensive English
program [IEP]) might differ in the quantity, quality, and/or variety
of strategies they used in several different settings, as well as how
background variables such as schooling and personality might
influence these strategies. (Other measures of success are, of course,
possible, e.g., ultimate level of achievement reached. However, since
theoretically, learners’ proficiency can increase throughout their
lifetimes, and since practically, we were interested in efficiency of
learning as well as level of proficiency attained, we used rate of
progress as our criterion of success.) In this paper, we briefly
describe the larger study, reviewing our earlier findings of one
successful and one unsuccessful learner as background for a detailed
discussion of two other learners from the unsuccessful group.

Our overall study examined 15 students, from a variety of lan-
guage backgrounds, in the intermediate to advanced levels of the
Iowa State Intensive English Program. At the time we selected
subjects, we did not know who would move quickly through the
program and receive passing TOEFL scores, and who would not.
However, after the subjects left the program, we ranked them
according to their rate of progress as measured by their average
weekly gains on the Michigan English Language Proficiency Test
based upon entrance and exit scores. The top four subjects, with
average weekly gains of 2.48 points (range = 1.69-3.13), we
defined as successful. The bottom four, with average gains of .75 points per week (range = .46-.98), we viewed as unsuccessful for the purposes of this study. These learners were not merely poor test takers; their instructors confirmed a similar lack of success in the IEP classes. The four successful and four unsuccessful learners were selected for closer analysis.

Along with test scores, and reports and questionnaires from teachers, we collected background information on these learners via interviews in which they told us about their experiences learning languages, and the strategies they used. We analyzed learners' interviews as well as their think-aloud discussions and performance on a series of four typical classroom tasks (a fill-in-the-blank verb exercise, a similar exercise on articles, a cloze passage, and a composition.) While these tasks resembled tests in format, the think-aloud technique encouraged negotiation of meaning between the subject and a research assistant, thus eliciting strategies believed to be associated with language learning.

METHOD

Think-Aloud Procedure and Analysis

We followed the procedures for gathering think-aloud data suggested by Hosenfeld (1976). Two research assistants, both Asian female graduate students in TESL with near-native command of English, collected the data. After an initial training session, subjects were asked to do tasks as they normally would, except they were instructed to “think aloud.” The two research assistants remained as unobtrusive as possible except to probe subjects’ thoughts if they were not being expressed, and to answer any questions subjects had concerning task procedures or the meaning of vocabulary. The research assistants were instructed not to provide other assistance. All sessions were audiotaped and transcribed. In analyzing the data, we identified and classified subjects’ strategies using a taxonomy developed by Rubin (1981), which we modified to more accurately reflect strategies actually appearing in our data. The major categories are shown in Figure 1.

Since data were collected under experimental conditions, results do not necessarily reflect what students would do under ordinary circumstances; that is, experimental conditions may have degraded or enhanced the number of strategies they used. However, we assume that strategies were not created for the experiments.

Phase I

Based on our own intuitions and most of the literature on
successful language learners, we predicted that the best learners would use more and a greater variety of strategies than the poorer learners. This expectation was supported by our initial analysis of two Spanish speakers, Gerardo and Pedro (all learners have been given pseudonyms), who contrasted dramatically in their rate of progress (Abraham & Vann, 1987).

Gerardo, one of the most successful language learners we had ever encountered, spent an enormous amount of time completing the tasks, using a large number of the strategies Rubin (1981) had

<table>
<thead>
<tr>
<th>STRATEGIES OF UNSUCCESSFUL LANGUAGE LEARNERS</th>
</tr>
</thead>
</table>
associated with successful language learning. In contrast, Pedro spent little time on tasks, used relatively few strategies, and had a smaller range of strategies than did Gerardo. It was not difficult to see why Gerardo succeeded in a formal language-learning situation while Pedro failed.

As we began to analyze the data from other learners, however, we discovered that the unambiguous contrast between Gerardo and Pedro was not typical; the number of strategies (as identified in our modified version of Rubin’s [1981] taxonomy) used by several of the successful and unsuccessful subjects fell within the same range (see Figure 2).

Some unsuccessful learners used relatively many strategies, and more important, the unsuccessful learners used many of the same strategies as the successful learners. This finding challenges the belief that unsuccessful learners lack strategies altogether or use ineffective strategies. Instead, it suggests that a broader range of strategies may be beneficial for all learners.

![Figure 2: Average Number of Strategies per Task Used by Most Successful and Least Successful Learners](image)

- **(F)** Successful
- **(Gerardo)** Successful
- **(Mona)** Unsuccessful
- **(Shida)** Unsuccessful
- **(Pedro)** Unsuccessful
- **(J)** Unsuccessful
- **(K)** Unsuccessful
- **(Pedro)** Unsuccessful
strategies as the successful learners. All learners, at least occasionally, clarified and verified, guessed at meaning, applied grammatical rules, monitored form, and demonstrated various production tricks and social management strategies. These findings called into question the common assumptions in our field, articulated by Wenden (1985), that “ineffective learners are inactive learners” and that “their apparent inability to learn is, in fact, due to their not having an appropriate repertoire of learning strategies” (p. 7). Our study suggested that, at least under these experimental conditions, unsuccessful learners were not necessarily inactive or lacking in their repertoire of strategies. What, then, was the cause of their failure? Simple strategy counts failed to provide the answer.

Phase II

The second phase of our study thus moved beyond the relatively straightforward images of success and failure we saw in Gerardo and Pedro to probe the more puzzling failures of two Arabic-speaking Saudi Arabian women, Mona and Shida. Though they had accompanied their husbands to the United States, they were modern Saudi Arabian women who were serious about pursuing their own schooling. Mona was a 20-year-old high school graduate, while 29-year-old Shida had completed college in Saudi Arabia. Both reported studying English throughout secondary school, but for only 3 to 5 hours weekly. Soon after they arrived, Mona scored 32 on the 100-item Michigan English Placement Test, while Shida scored 41 on the same test, thus placing as high beginners in the intensive English program in which they were enrolled full time. After 40 weeks in the program, Mona achieved only a 430 on the TOEFL; after 24 weeks, Shida scored only 413. (Mona and Shida scored 150 and 170 respectively out of a possible 300 on the Test of Spoken English.) Unlike the unsuccessful Pedro, these women had relatively high strategy counts, yet failed in spite of doing what appeared, at least superficially, to be “all the right things.”

To gain insight into this problem, we reexamined the data from a new perspective, this time beginning with the tasks and the demands they made on learners. We then assessed Mona’s and Shida’s strategies to determine their effectiveness in meeting these demands. For this analysis, the most revealing data proved to be in the interviews and in responses to three of the four tasks: the verb exercise, the cloze passage, and the composition. Demands for successfully accomplishing these can be classified within four categories: engagement, risk taking, knowledge, and control.

All tasks in this study required engagement, a factor identified by
Jakobovits (1970) as critical to language learning. (This factor is similar to “attention” as discussed by Scovel, 1989.) Here engagement meant spending sufficient time on the assignment, clarifying and verifying the task demands where necessary, and providing evidence of attentiveness. Both Mona and Shida were engaged in all tasks they performed for this study.

The tasks also required what Beebe (1983) has called risk taking. One can argue that all second language learning requires learners to take risks, but situational variables and learner characteristics can render some tasks more risky than others for a given learner. Learners in this study took substantial risks when they attempted tasks or items for which they did not know the procedures or the answers (such as completing a cloze passage or, in the case of the composition task, writing a police report).

Knowledge demands for the tasks were of three types. First was procedural knowledge, or knowing how to proceed on a task. Second was background knowledge, what theorists have termed schemata, which would enable learners to fit the new information presented in the task into their already established framework of knowledge. (For an extended explanation of schemata, see Rumelhart & Ortony, 1977.) The last type was knowledge of the language itself. As noted by Bialystok and Ryan (1985), this knowledge could be either analyzed (i.e., a self-conscious awareness of features and functions, making it possible, for example, to articulate a grammatical transformation) or unanalyzed (i.e., intuitive and unconscious knowledge, evident in statements like “I don’t know why it is correct, but it sounds right”).

Finally, the tasks demanded varying levels of cognitive control, processes that manage selection and coordination of knowledge. As Bialystok and Ryan (1985) note, control in coordinating information becomes increasingly important where monitoring procedures are needed to oversee several aspects of a problem, for example, form and meaning, or meaning and context.

It is important to emphasize that these factors are primarily task dependent, not learner dependent. Our tasks were arranged so that they progressed from least to most demanding along the dimensions of engagement, risk taking, knowledge, and control. However, it should be noted that these factors sometimes intersect; for example, insufficient knowledge for a task may cause a learner not to engage.

RESULTS: TASK DEMANDS AND PERFORMANCE

Activity 1: Interview

Task Demands. This informal situation, which was basically conversation with an empathetic interviewer, made relatively low
demands, at least on subjects who were willing to risk speaking in English.

**Performance.** Both Mona and Shida were engaged and cooperative. Both used a number of strategies that enhanced communication (e.g., clarification/verification of meaning, comprehension checks). Interestingly, both frequently monitored their pronunciation, vocabulary, and grammar. While they often produced ungrammatical sentences, they usually managed to convey their ideas, thus meeting the demands of the task.

**Activity 2: Verb Exercise**

The verb tense task was a standard, sentence-level classroom exercise requiring the student to fill in blanks with correct verb forms (either simple present or present progressive). It contained 16 items (see Appendix A).

**Task demands.** This task demanded relatively low engagement and risk taking. Procedural knowledge demands were also low, since the task type was familiar to the students. Likewise, little background knowledge was required, since the items involved isolated sentences, often with adverbial of time and frequency as clues. Linguistically, the task required only that subjects distinguish between contexts requiring simple present and present progressive tenses and then to correctly form each tense. With respect to control, demands were also fairly low. Subjects had to recall the appropriate rules accurately and apply all parts of the rules correctly (for example, BE + verb + ing).

**Performance.** Mona’s score was 15 correct out of 16, the second highest for all subjects, while Shida scored 11 out of 16, the third from lowest. Mona’s high score is consistent with both adequate knowledge and control for this task. However, Shida, while achieving moderate success with a “sounds right” strategy (“I don’t know what to say, I think it’s good”), demonstrated problems in applying conscious rules. Although she attended to some formal aspects of her response (e.g., changing y to i in tries), she was erratic in her application of the BE + verb + ing rule (once using it correctly, twice not). She did not choose to check her answers after she finished, thus confirming the view that she lacked a systematic approach to monitoring her work.

**Activity 3: Cloze**

The cloze exercise (reproduced with acceptable responses in Appendix B) was constructed from a passage on cross-cultural differences in the use of time and space, and contained 33 blanks.
As expected, performance on the cloze was, in general, closely related to overall proficiency. Since both Mona’s and Shida’s overall proficiency was low, the cloze proved difficult for them.

**Task Demands.** Compared to Task 1, this activity required considerably more knowledge and control, as well as more engagement and willingness to take risks. The procedural knowledge required for the cloze passage was greater because, as Mona noted, unlike most classroom fill-in-the-blank tasks, the student here received no guidelines delimiting possible answers (as in Task 1).

To fill the blanks correctly, subjects needed considerable knowledge about cross-cultural differences in proxemics and attitudes toward time. This could be obtained from the passage itself, but for students of low proficiency, reading a mutilated passage for meaning was not easy.

As for language-specific knowledge, the task required subjects to recognize and produce vocabulary relevant to the topic, and to know enough about grammatical structure and rhetorical conventions to choose words that created well-formed sentences and sensible text.

With the different kinds of knowledge required, this task placed considerable demand on control, involving constant coordination of structure and meaning.

**Performance.** We scored this task crediting all acceptable responses. Mona, with 8 out of the 33 items correct, was the third from lowest for all subjects; Shida, with 10 out of 33, the fourth from lowest. Although these scores were similar, comments in the think-aloud data and analysis of the written responses reveal striking contrasts between the two learners.

Mona’s think-aloud data indicate that she used two major strategies in filling the blanks. The first she articulated as “[making] something similar,” i.e., repeating a word that had occurred in a similar phrase earlier in the passage. Sometimes this strategy resulted in a correct answer, as in blank #1, “In most parts of Latin America, thirteen inches is just the right distance when talking.” More often it did not, as in blanks #10 and #11, “All they know is that inches is something wrong with the other distance.”

Mona’s second major strategy, used for the majority of blanks, was to look for words that were compatible with the words or phrases immediately surrounding the blank. Again, this produced a few correct answers, as in #13, “Most culture-blind Americans think that Latins are pushy,” but usually it failed because she did not take into account enough of the text. For example, she used about instead of to, to fill blank #2 (“thirteen inches is just the right...}
distance when talking ______ a person”) because she took the necessary context to be “talking ______ a person,” noting that when one talks, it is usually about something. Other examples were #6, “The Latin will be to maintain,” and #14, “In most urban areas, will be two minutes [late] for an appointment is all right,” because “usually it’s will be.” Throughout the exercise Mona relied heavily on this localized approach, frequently plugging in bits and pieces of formulaic expressions and rules that she had learned in English classes or conversation. Conscious of form as she worked, she also checked her answers after she had finished, changing two, but once again, using local cues. With respect to obtaining the overall meaning of the passage, her think-aloud data indicate that she did not consider this an important element of the task.

Shida differed in her approach. While she also used local strategies, she was very much concerned with trying to get the gist of the passage. Upon encountering difficulties a quarter of the way through the exercise, she went back to reread the beginning and, identifying a word (stand) that she apparently felt represented a key concept, asked the research assistant for its meaning. She repeated this strategy of asking for meaning eight more times in attempting to come to grips with the passage, contrasting sharply with Mona, who asked for only one definition.

Some of Shida’s answers also contrasted with Mona’s, reflecting the former’s search for meaning. For #10 and #11, she produced, “All they know is that possible (semantically appropriate) is something wrong with the other side (acceptable).” Again, for #6, she wrote, “The Latin will chose (semantically correct) to maintain . . . .” The problem with each of the incorrect answers above is one of form, resulting from either a lack of knowledge or failure to retrieve what she had learned earlier. Once again, though she monitored regularly as she worked (particularly in correcting reading miscues and faulty pronunciation), Shida opted not to check her answers after she had finished.

One final contrast between Shida and Mona is worth noting. Shida worked steadily to fill the blanks as best she could, although she made it clear that she knew her finished product contained errors. Mona, on the other hand, remarked on the difficulty of the task early on (“It’s kind of hard one”). She then pleaded with the research assistant for help, asking her to shake her head when Mona’s answer was incorrect: “You can help me. If you won’t help me, I can’t continue because I’m not sure if I’m wrong.” After an apology (“Really, I need more practice”), in frustration, she tried to alter the rules: “Give it [the assignment] to me and I promise I’ll bring it to you tomorrow.” Seeking help from the assistant, rather
than relying on either her own resources or a text, fits Mona’s highly interactive style, a personality trait that handicapped her on many written tasks. A similar phenomenon has been observed in children (Bereiter & Scardamalia, 1987): With their preference for the interactive mode (designed to respond to signals from the external environment, specifically from conversation partners), children find writing tasks (which are divorced from this interaction) especially challenging.

In summary, although neither subject had sufficient knowledge to excel in this task, their approaches to managing the situation differed, probably reflecting their assessment of the task demands. Mona, when faced with a task beyond her ability, employed low-level strategies (careful application of grammatical knowledge and local context cues) similar to those she had employed successfully in the verb-tense exercise. Significantly, when faced with difficulties, she attempted to change the rules to accommodate her capabilities rather than to alter her procedure. She attributed her difficulties to a lack of practice in doing this kind of exercise. In contrast, Shida, while using some of these low-level strategies, also used appropriate higher-level strategies, in which she looked for overall meaning and then forms to convey these. The flaw in her approach was in not attending sufficiently to form after she had established the meaning.

Activity 4: Composition

This was an unusual composition assignment, requiring subjects to write a police report of a traffic accident (see Appendix C).

Task Demands. Since the format of the composition was unfamiliar to all of the subjects, forcing them to attempt something they had never done before, engagement and risk taking were important in this task. Subjects needed the procedural knowledge to read a diagram and to set up a police report. They also needed background knowledge concerning the police report schema, that is, knowing what kinds of information are relevant to this genre. The required linguistic knowledge included vocabulary to describe the accident, syntax to write well-formed sentences, discourse conventions for making a report, and the mechanics (punctuation, capitalization, etc.) used in formal writing. With respect to control, subjects needed to select and coordinate appropriate information from knowledge stores.

Performance. Four independent ratings of the composition products for all subjects were averaged. Of the 15 students, Mona was rated 12th and Shida 13th, both in the bottom 20%.

Like the successful learners, Mona and Shida devoted attention to dealing with their initially inadequate knowledge for this task by
(a) clarifying task demands, (b) identifying the meaning of key vocabulary, and (c) understanding content (what happened in the accident they were to report). Both Mona and Shida were also engaged in acquiring the schemata and procedural knowledge for completing the task. (It is highly improbable that either had ever seen an accident report written by a police officer, and even less likely that they had been asked to do such writing.) Yet this was a source of frustration only for Mona, who asked again and again what was expected, compared the task before her with previous writing assignments, and was troubled that formulae used successfully in them did not apply here. In other words, she did not take the risks that this task demanded. While both learners spent sufficient time on the task (as compared to the time spent by others), Shida used this time more efficiently by settling discrepancies between what she saw in the text and what she knew to be true about accidents (establishing her schemata), and employed strategies that have been associated with expert learners, e.g., determining how to handle what she considered insufficient data—no license numbers were given. (For discussion of experts vs. novices, see Bransford, Nitsch, & Franks, 1977). She entered the task, adding her own interpretation of who was at fault in the accident and joking about adding her signature at the bottom of the report, evidence of her ability to deal with an assignment that did not conform to typical classroom experience.

The key difference in behaviors between the two learners on this task was that, for Mona, a good composition was one that was neatly written and grammatically correct, a matter of formulae plus practice. Mona’s handwriting was clear and legible, and her insistence on recopying the exercise testifies to the importance she assigned to neatness. Similarly, while Mona’s composition was not free of mechanical errors (both she and Shida had errors in spelling, capitalization, and punctuation), it was superior to Shida’s in this regard. At the grammatical level, Mona’s report was uneven, on the one hand containing an incorrect form of the simple past ("I was saw an accidents yesterday"), but on the other hand containing a correct form of the past perfect. Interestingly, in adding a relative clause construction to her revised version ("Than the Pikup went outside Northbound to Southbound, which passed the Chevrolet station with skip marks going from the Northbound to Southbound and to Northbound"), Mona sacrificed the clarity of this segment in her first draft. Once again, Mona’s monitoring strategies focused on form rather than meaning.

Shida’s work, with its sloppy script and numerous mechanical and grammatical errors, nevertheless was clear and coherent in meaning
throughout. Her report contained only simple past tense verb forms, which she overgeneralized. She attempted no embedded sentences, but simply connected ideas with and throughout as a means of achieving a connected text. Unlike Mona, Shida added her own interpretation of the event: “The driver of a pickup truck made a big accident, so this accident only his fault, and I think he must pay a fine and be punished, because he made a big mistake.”

Comments from the two learners during the think-aloud procedure reinforce the picture we gained from examining the learners’ products, Mona saw this assignment as a linear problem with a single correct solution. In contrast, Shida viewed the exercise more broadly. For her, understanding of meaning was fundamental, as witnessed in her attempts to understand the task demands and the content required, reconcile the information in the diagram with the written information presented, and become personally involved in the assignment by taking on the role of the police officer.

In summary, Mona’s approach to composing was superficial and formulaic, and her monitoring strategies focused on form. In contrast, Shida used the meaning-based control strategies associated with expert composers, who consider “. . . whether the text they have written says what they want it to say and whether they themselves believe what the text says” (Bereiter & Scardamalia, 1987, p. 11). Yet, despite her use of this strategy associated with mature writers, she lacked, or at least failed to apply, a systematic approach to checking form.

DISCUSSION

In this study, we have gone beyond collecting generalized self-reports to develop a model of analysis that incorporates both think-aloud protocols and task products. Rather than merely tallying various types of strategies, we have linked strategies with task demands. This procedure has permitted us to pinpoint discrepancies between what a given task requires and what the learner does. Where the noncontextualized strategy counts failed to differentiate unsuccessful from successful learners, or, for that matter, unsuccessful learners from each other, this model permits considerable insight into strategies used by language learners.

The microanalysis of Mona and Shida provides counterevidence for the claim that unsuccessful learners are “inactive” (Wenden, 1985, p. 7). In fact, when using our earlier taxonomy of strategies to compare Mona and Shida with successful learners, we found these two subjects to be remarkably similar to successful learners in their repertoire of strategies. When viewed through the lens of the task-demand model proposed in this study, Mona and Shida still appear
to be active strategy-users, but they often failed to apply strategies appropriately to the task at hand. Apparently, they lacked certain necessary higher-order processes, what are often called metacognitive strategies or self-regulatory skills (see Wenden, 1987b), which would enable them to assess the task and bring to bear the necessary strategies for its completion.

Mona’s careful application of rules and monitoring of errors served her well on relatively simple tasks but became a liability on more complex ones. It may be that, as Bereiter and Scardamalia (1987) found for novice writers, the attention she needed to devote to one aspect of a difficult task left her little spare capacity for implementing more sophisticated executive procedures. According to this explanation, as Mona gained control over various subprocesses and acquired further proficiency, she would be able to devote more attention to the appropriate regulatory skills. An alternative explanation, made more plausible by her slow progress, is that Mona’s reliance on formulaic approaches to higher-level tasks may have been promoted by earlier formal instruction, which either explicitly or implicitly encouraged her to seek straightforward explanations to problems, however complex. In Bereiter and Scardamalia’s (1987) terms, her schooling may have encouraged “knowledge-telling” rather than “knowledge-processing” (pp. 5-6).

In contrast, Shida’s strength lay in pursuing meaning, but she applied no systematic set of strategies for attending to form. As with Mona, this may have been a matter of cognitive overload, but, interestingly, Shida chose the higher-level procedure while forfeiting attention to detail. Another explanation is that Shida was applying strategies from oral communication, her forte, concluding that these “little details” were relatively unimportant (as in oral language).

The question remains: Would strategy training benefit learners like Mona and Shida? If learners lack a strategy altogether (which Mona and Shida did not), or if they do not appropriately link the strategy of which they have command with the task before them, then training may be appropriate. However, to make a strategy generally applicable, learners apparently need to be convinced of its significance and be taught to evaluate its use (see review in Wenden, 1987b). Brown, Day, and Jones (1983) note that strategy training studies providing specific information about why, when, and where strategies should be applied have been the most successful in inducing transfer.

The insights of this study raise further questions: Are the approaches of learners such as Mona and Shida relatively stable or variable? Do these approaches fall into predictable patterns when
extended to larger groups of learners? How are learning approaches related to learner variables such as personality, age, and formal schooling? Finally, how useful is strategy training in altering what appear to be counterproductive learning approaches? Longitudinal studies that follow learners over extended periods of time, work that examines large groups of learners, and research that considers learner variables in relation to metacognitive strategies would make important contributions to our knowledge about learner strategies.

Continued case studies are essential in advancing this area of study. That the two apparently similar language learners in this study differed significantly in their fundamental approaches to problem solving underscores the insights provided by microanalysis of learner behavior on varied tasks. Future case studies will enable researchers to verify critical assumptions about second language learning.

ACKNOWLEDGMENTS

An earlier version of this paper was presented at the 23rd Annual TESOL Convention in San Antonio, TX, March 1989. The authors wish to acknowledge the cooperation of the Intensive English and Orientation Program at Iowa State University, and support for this research granted them by Iowa State University. They also wish to thank Andrew Cohen, Joan Rubin, and two anonymous TESOL Quarterly reviewers for helpful comments on an earlier draft of this article.

THE AUTHORS

Roberta J. Vann is Associate Professor of English at Iowa State University, where she teaches in the Master’s program in TESL. She is a previous contributor to the TESOL Quarterly and coauthor of Viewpoints USA (Harper & Row, 1984).

Roberta G. Abraham is Associate Professor of English at Iowa State University, where she teaches courses in TESL and linguistics. Her research has focused on the relationship between learner characteristics and success in language learning.

REFERENCES


### SIMPLE PRESENT TENSE vs. PRESENT PROGRESSIVE TENSE

<table>
<thead>
<tr>
<th>Simple present</th>
<th>Present progressive</th>
</tr>
</thead>
</table>
| 1. expresses repeated action (includes the past, present and future) | 1. expresses one action in the present 
   a. of short duration 
   He's writing a letter. |
| The earth revolves around the sun. (general truth) | b. of long duration 
   He's studying English. |
| I go there very often. (custom) | He's writing a book. |
| 2. expresses non-action (state or condition) | 2. expresses future action |
| He seems tired. | He's giving a lecture tomorrow. |
| She loves her children. | The ship is sailing next week. |
| I remember him. | |
| I hear some music. | |
| (vs. I am listening to some music.) | |
| 3. expresses future action (especially with verbs of arriving and departing) | 3. expresses the beginning, progression or end of an action |
| We leave tomorrow. | It is beginning to snow. |
| The ship sails next week. | My cold is becoming worse. |

Supply the simple present or the present progressive form of the verb. In a few sentences either form may be used.

**EXAMPLE:**

a. The milk (taste) ________ tastes ________ sour.
   b. She (taste) ________ is tasting ________ the soup to see if it needs more salt.
   c. The wind (blow) ________ is blowing ________ very hard outside.

1. The play (begin) ________________ now.
2. She (try) ________________ to finish her work early today.
3. It (get) ________________ colder and colder.
4. I (hope) ________________ to see you again.
5. We (plan) ________________ to buy a house soon.
6. Children (learn) ________________ faster when they are interested in what they (study) ________________.
7. We (go) ________________ to the movies tonight.
8. The sun (rise) ________________ in the east and (set) ________________ in the west.
9. I sometimes (forget) ________________ to take my keys when I (leave) ________________ the house.
10. She (take) ________________ a nap every afternoon.
11. I (hear) ________________ some loud noises outside.
12. He (listen) ________________ to the radio.
13. I (see) ________________ some children outside.

INSTRUCTIONS
In the following passage 100 words have been omitted. Read the passage and insert whatever
word makes sense according to the meaning of the passage. The word should be
grammatically correct. Remember—insert only ONE word in each space. Read the whole
passage at least once before you start to write.

Example: The boy walked ___ across the street and bumped ___ into ___ a lamppost. He
___ was ___ shaken up a little, but he managed to ___ continue ___ walking.

The Jet Age Malady
A U.S. male brought up on the East Coast of America stands eighteen to twenty inches
away from another male when in conversation. In talking to a woman he will increase the
distance by about four inches. To stand at a distance of about thirteen inches usually has a
sexual or aggressive connotation. However, in most parts of Latin America thirteen
inches ___ is just the right distance when talking ___ with, to ___ a person.

When a man brought up in a ___ Latin, South ___ American environment tries to talk to a
person, man ___ brought up on the East Coast of ___ the ___ United States an
interesting thing happens. The Latin will ___ try, attempt ___ to maintain what he considers
the ___ right, proper ___ talking distance. The American will, of course, step
back, aside ___ Both will feel uncomfortable without quite ___ knowing, realizing ___ why.

All they know is that ___ there ___ is something wrong with the other
person, side ___ Most culture-blind Latins feel that the Americans ___ are ___ withdrawn and uncommunicative. Most culture-blind Americans ___ feel, think ___ that
Latins are pushy.

In most American urban areas, ___ to ___ be two minutes ___ late ___ for an appointment is all right. Three ___ minutes ___ is significant, but an apology is not
expected. After five minutes the latecomer mutters an apology. In Latin countries a five-minute unit is not significant, important: an apology is expected only for a time period, much longer than twenty minutes. Latins, influenced by their own cultural conditioning, feel that Americans are not polite and are obsessed with time because they expect persons with whom they have appointments to be at a certain place at precisely a certain, given time. A person unfamiliar with North American cultural conditioning has difficulty realizing that Americans handle time much like some tangible material—spending it, taking it, using it up, or wasting it. If, When a Spanish-American or a Spaniard comes to work late he says, “El bus me dejó” (“The bus missed, left me”), as opposed to the American “I missed the bus.” In English, the clock “runs.”

---


Note: Acceptable responses are indicated in italics above each blank.
You are a traffic officer. As part of your job, you must file a report of accidents you covered while on duty. Yesterday, you were on the scene of an auto accident that took place on a country road. You now need to file a report of that accident.

Task
Write a report of the accident. The following information is what you wrote down in your note pad. Use this information to write your report on what happened yesterday. Be certain to make clear the sequence of events.

Time: 7:20 A. M., April 14
Place: Highway 652, two miles south of the city
   An overturned Volkswagen on the shoulder of the southbound lane
   Skid marks leading from the southbound lane to the Volkswagen
   A pickup truck blocking the northbound lane of traffic
   Skid marks going from the southbound lane into the northbound lane
     (leading to the pickup truck)
   Front of a Chevrolet station wagon smashed against the side of the pickup truck

The Discourse of Computer-Assisted Language Learning: Toward a Context for Descriptive Research

CAROL CHAPELLE
Iowa State University

Understanding how the speed, power, and flexibility of computers can facilitate second language acquisition is an intriguing challenge faced by instructors, researchers, and theorists. Progress in this area, however, does not appear to be forthcoming from current research on computer-assisted language learning (CALL), which suffers from the same limitations as early research on classroom instruction: Little detail is provided to describe the interaction among participants during instruction (Long, 1980). Moreover, descriptions of CALL activities included in reported research are not empirically based: They fail to describe what subjects actually do while working with CALL. A third problem is that the terms used to describe CALL activities have been developed specifically for that purpose, and are therefore not comparable to those used for classroom activities. At the same time, these descriptors are not sufficiently uniform and formally stated to allow specific comparisons among CALL activities. Toward a solution to these problems, this paper proposes a discourse analysis of student-computer interaction enabled by viewing the student and the computer as two participants in a dialogue. It argues that the discourse analysis system of classroom interaction developed by Sinclair and Coulthard (1975) provides the necessary elements and structures to describe CALL discourse, analyze data from student-computer interaction, and compare CALL activities with other (classroom) activities.

Computer-assisted language learning (CALL) is now used routinely in language instruction (e.g., to provide out-of-class practice in grammar and reading skills, problem solving, group work, and writing); as technical capabilities and human imaginations expand, additional uses for computers in ESL instruction will emerge. Instructors and researchers need to understand how CALL
can best be used to offer effective instruction to language students. More than other resources, CALL has the potential for individualizing instruction. Accordingly, a CALL research agenda should seek concrete results concerning successes and failures of individual students with a variety of CALL activities. Unfortunately, the lack of precision characterizing CALL research at present precludes such results. Little if any current CALL research can offer unambiguous evidence concerning effects of CALL activities because current research methods fail to elucidate exactly what students do while they work with language learning software.

Much current CALL research (e.g., Kleinmann, 1987) shares the pitfalls of investigations of second language classroom teaching methods of the 1950s and 1960s, in which performance of students in classrooms labeled, for example, grammar-translation was compared to students' performance in classrooms labeled, for example, audiolingual. (See Allwright, 1988, for a detailed description of this history.) As Long (1980), Allwright (1988), and Chaudron (1988), among others, point out, this research was inconclusive because too many factors influencing students' performance were not accounted for. One of these factors was what students and teachers actually did and said in the classrooms under investigation. To underscore the importance of describing classroom interaction, Long (1980) noted that "there is no guarantee . . . that the methods do not overlap in some respects, e.g., through their shared use of a common subset of classroom procedures. In a research context, this means that it is impossible to ascertain which subjects have received the treatment" (p. 2). Classroom research requires more than general labels for instruction; it requires precise descriptions of the interaction that occurs in classrooms. Similarly, if researchers hope to understand what and how particular students learn using CALL materials, it is necessary to characterize the interaction that takes place while they work.

This paper clarifies the need for a precise analysis of student-computer interaction in CALL research, using the principles of classroom discourse analysis developed by Sinclair and Coulthard (1975). (All subsequent references to Sinclair and Coulthard cite this 1975 description.) The paper illustrates how such an approach might be used to characterize systematically students' work on a CALL grammar lesson. This discourse analysis approach enables a description of potential and actual student-computer interaction as students work with the lesson, and provides a comparison of the grammar lesson with other activities. We begin with a summary of current approaches for describing CALL activities, noting their shortcomings.
CURRENT APPROACHES TO DESCRIBING CALL

The CALL literature abounds with schemes for describing and categorizing the computer activities used in language classes. In pedagogical terms, *program-controlled* vs. *learner-controlled* activities are dichotomized. Making this distinction, Higgins' (1988) terms "magister" (p. 12) and "pedagogue" (p. 14) provide a vivid analogy by anthropomorphizing the two types of programs. The magister is the powerful instructor in control of the students and what they learn. The pedagogue role of the computer is that of helper; the student is in charge of learning and bears the responsibility for calling on the computer as needed. These definitions (detailed in Figure 1) of the roles of computers in language instruction have underscored the fact that there is nothing inherent in computers to render their role in the classroom magisterial.

FIGURE 1
Higgins' (1988) Parameters for Describing CALL

<table>
<thead>
<tr>
<th>Pedagogical parameters</th>
<th>Magister</th>
<th>Pedagogue</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Directs students' learning</td>
<td>Assists students' learning</td>
</tr>
<tr>
<td></td>
<td>Has authority to evaluate, praise, censure</td>
<td>Has no authority</td>
</tr>
<tr>
<td></td>
<td>Structures the order of events</td>
<td>Provides no structure</td>
</tr>
<tr>
<td></td>
<td>Explains rules; gives examples</td>
<td>Answers students' questions</td>
</tr>
<tr>
<td></td>
<td>Repeats endlessly</td>
<td>Follows students' orders</td>
</tr>
</tbody>
</table>

The magister/pedagogue distinction, however, fails to account for linguistic aspects of computer activities, that is, the types of language that are the focus of an activity. This linguistic dimension is what Underwood's (1984) dichotomy for computer activities

---

1 Other terms used for *program-controlled* and *magister* are “knower-of-the-answer” referring to tutoring activities (Jones & Fortescue, 1987, p. 5), “tutor” (Kenning & Kenning, 1983, pp. 2-3), and “instructor” (Wyatt, 1984, p. 6).

2 Other terms consistent with *learner-controlled* and *pedagogue* are Jones and Fortescue’s (1987, p. 6) “knower-of-the-answer” (when the computer is used as a discovery device), “workhorse,” and “stimulus”; also Wyatt’s (1984, p. 8) “facilitator” (when the computer is used for wordprocessing or data bases) or Wyatt’s (1984, p. 7) “collaborator” (when the computer is used for adventure games and simulations, for example).
offers by distinguishing “communicative” (p. 51) from noncommunicative CALL. (Note that Underwood’s term for the opposite of “communicative” software is “wrong-try-again” [p. 52] software rather than “noncommunicative” as used here.) Underwood’s “premises for communicative CALL” (p. 51) identify features (summarized in Figure 2) he presumes communicative CALL to have. Underwood’s premises are based on Krashen’s (1982) input hypothesis, which assumes that comprehensible input is what is needed to allow the learner to develop an acquired linguistic system; explicit instruction, therefore, is of very little benefit because such instruction serves only the “learned” system. Problems with this account of second language acquisition are outlined by McLaughlin (1987). Despite the questionable theoretical constructs forming the basis for these features, they succeed in providing criteria for judging whether or not a computer activity is communicative. Moreover, Underwood’s criteria for communicative CALL have helped to emphasize the importance of meaningful language use in computer activities.

FIGURE 2
Underwood’s (1984) Parameters for Describing CALL

<table>
<thead>
<tr>
<th>Linguistic parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Noncommunicative CALL</strong></td>
</tr>
<tr>
<td>Program incorporates grammatical sequencing</td>
</tr>
<tr>
<td>Program judges to inform students of their errors</td>
</tr>
<tr>
<td>Program is in control</td>
</tr>
<tr>
<td>Subject matter is irrelevant to student</td>
</tr>
<tr>
<td>Lesson is predetermined</td>
</tr>
<tr>
<td>Student perceives task as a required lesson</td>
</tr>
<tr>
<td>Student views task as identical to classroom activities</td>
</tr>
</tbody>
</table>

| **Communicative CALL**                 |
| Program does not impose grammatical sequencing |
| Program judges more to provide helpful hints |
| Student is in control                   |
| Student relates to subject matter in a personal way |
| Student creates own learning experience |
| Student perceives task as motivating supplement |
| Student views task as a novel activity |

Despite the values of these general pedagogical and linguistic software definitions, their lack of precision is apparent to anyone who has attempted to use them to write a detailed software description. Seeking greater precision in describing CALL activities,
Phillips (1985) provides a set of descriptors integrating pedagogical and linguistic categories. These terms (listed in Figure 3) can be used both to draw informal distinctions among kinds of software and to generate ideas concerning possibilities for developing CALL activities. Phillips' terms, such as “quiz activity,” (p. 26) are designed to suggest what students might actually be doing in a computer activity, just as the name communicative classroom, for example, brings to mind particular classroom activities. Phillips' description of a rational-deletion cloze CALL exercise is summarized in Figure 3, using his complete set of parameters. Note that this descriptive scheme still fails to capture the actual activities of students engaged in CALL.

### Figure 3
Phillips' (1985) Description of a Rational-Deletion Cloze CALL Exercise

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Cloze exercise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity type</td>
<td>Quiz</td>
</tr>
<tr>
<td>Learning style</td>
<td>Recall, experimental learning</td>
</tr>
<tr>
<td>Learner focus</td>
<td>Doing a test</td>
</tr>
<tr>
<td>Program focus</td>
<td>Control over syntactic form</td>
</tr>
<tr>
<td>Classroom management</td>
<td>Individual work</td>
</tr>
<tr>
<td>Program difficulty</td>
<td>Little flexibility</td>
</tr>
<tr>
<td>Language difficulty</td>
<td>Could be a choice of levels</td>
</tr>
</tbody>
</table>

Results of CALL research are reported in terms similar to Phillips'. The following exemplify the type of descriptors used: “text manipulation programs” (Piper, 1986, p. 187); “simulation” (Jones, 1986, p. 179); “drill and practice,” “free conversation,” “problem-solving simulation” (Abraham & Lieu, in press); “software packages emphasizing reading comprehension skills” (Kleinmann, 1987, p. 269); “a reading skills program which automatically provides cloze versions of texts” (Windeatt, 1986, p. 82); “drill and practice . . . grammar, reading, and listening” (Chapelle & Jamieson, 1986, p. 30). Unfortunately, these terms are as inadequate for describing activities under investigation in CALL research as are the names of classroom methods for the precision required of successful classroom research. Consequently, it is rarely

3 Pedagogical approaches used in courseware have been detailed by educators for years, using such terms as drill and practice and simulation (e.g., Allessi & Trollip, 1985). The linguistic dimension has been categorized using the traditional skill areas such as reading and grammar (e.g., Wyatt, 1984; Ahmad, Corbett, Rogers, & Sussex, 1985).  

THE DISCOURSE OF COMPUTER-ASSISTED LANGUAGE LEARNING
clear exactly what students have done while they were working on the CALL activity under investigation. This lack of precision makes CALL research to date difficult to interpret. Three problems arise in conducting and interpreting CALL research when instructional activities are described solely by the researcher’s holistic impressions.

First, descriptions based on a single view of an entire CALL activity do not account for the details of student-computer interaction. In the “quiz” activity cited above, for example, left unspecified were pedagogical features including the following: the order in which the blanks are completed, the number of opportunities given the student to complete a blank, and the help and exit options available. The unspecified linguistic details include types of deleted words, specific student errors, and the computer’s assessment of and response to those errors.

A second problem with general descriptions of CALL activities is that they characterize what students can or should do while working on a computer activity, failing to describe what students actually do. Simply labeling a CALL program a “quiz activity” does not necessarily make it one. If the student uses the “quiz” to “explore the limits of the computer’s ‘knowledge’ by devising suitable examples to test out the rules it is using” (Phillips, 1985, p. 29) rather than to try to get the correct answers on the first try so as to accumulate a high score, then that student’s behavior would be inconsistent with the program’s designation as a quiz. Instead, the student would be treating the quiz as an exploratory program. Similarly, with respect to linguistic strategies, if the program developer intends a cloze activity to allow students to use discourse clues (clues beyond the sentence) to develop discourse competence, yet the student consistently guesses on the basis of the immediate context (i.e., within the clause), then the developer’s program description would again be inaccurate.

These two inadequacies in descriptions of CALL activities negatively affect the internal validity of CALL research because incomplete descriptions of the instructional materials under investigation preclude unambiguous results. Moreover, misconceptions concerning what subjects did during an instructional activity ensure inaccurate interpretation of results.

The third problem affects the external validity, or generalizability, of CALL research. When a CALL activity is described in terms devised and defined exclusively for that exercise, it is not clear how the exercise is similar to or different from other CALL or classroom activities. To date, descriptors have been used to establish pedagogical categories under which CALL activities can be labeled, but have not yet specified a precise language (or
formalism) for discussing different CALL activities in terms common to all CALL and classroom activities. Such general descriptions do not allow the researcher to address the following: How quiz-like is a quiz? What additional features would turn a quiz into a drill? What features does a CALL quiz share with a classroom quiz? How is a CALL quiz different? What language forms and functions does the activity require of the student? The linguistic descriptors typically used for CALL are far more structure-oriented (e.g., lexis, morphology) and skill-oriented (e.g., reading, writing) than are the functional communicative terms used by classroom researchers. (Functionally defined terms used in current research on second language classrooms are illustrated in collections by Larsen-Freeman, 1980; Day, 1986; and Fine, 1988.) For valid comparisons of the capabilities and limitations of CALL relative to classroom activities, it is necessary ultimately to describe CALL interaction using terms similar to those used by classroom researchers.

In summary, identifying the effects of CALL activities on specific learners requires precise description of the interaction (or discourse) that occurs between learner and computer. Such a description must be empirically based, and expressed in a formalism that can be used for all CALL, as well as classroom, activities. Such a descriptive formalism has been explored by classroom researchers who have recognized these problems and have been attempting their solution for the past 15 years. This paper argues that the classroom discourse analysis system of Sinclair and Coulthard provides a promising direction for empirically based, formal descriptions of CALL.

DESCRIPTION OF CLASSROOM DISCOURSE

Researchers of language classrooms hope ultimately to understand the effects of classroom instruction by first systematically describing what teachers and students do and say in classrooms. In second language research, for example, Chaudron (1977) employed elements of the classroom discourse analysis of Sinclair and Coulthard as the basis for his description of teachers’ corrective treatment of language learners’ errors, using some of Sinclair and Coulthard’s functional units and their sequencing rules. (See Chaudron, 1988, for a comprehensive discussion of units of analysis in second language classroom research.) Sinclair and Coulthard attempt a thorough discourse description of the kinds of acts found in classrooms and the constraints on their location in a lesson as a whole. Because of the context dependency of functional definitions of discourse units, they place discourse acts within a larger structure, thereby describing the sequencing of particular acts.
This system (simplified and adapted here), based on observation of language used in classrooms, proposes a category system consisting of analytic units related to each other in a structure called a “rank scale” (p. 20). Higher ranks are defined by the units of lower ranks they comprise. Devised to build a hierarchical description of the acts composing classroom lessons, the system has at its lowest rank the acts that teachers and students perform. An act might be something like the teacher’s question, “Does anyone know?” That particular act, according to the linguistic rules, could be used as an initiating move, the next higher unit. There are several different types of moves in the original system, three of which are initiation, response, and feedback (also termed follow-up). Teachers often, but not always, initiate when they ask students questions. Students usually respond. Teachers often follow up when they comment on, evaluate, or elaborate on students’ responses. When those three moves together occur in a sequence, they are further analyzed as an exchange. Sinclair and Coulthard defined the sequence (initiate, respond, feedback) as a particular type of exchange—a teaching exchange. At the next level then, a series of an unspecified number of teaching and other types of exchanges compose a transaction. A series of transactions is, in turn, analyzed as a lesson. (Note that although Sinclair & Coulthard’s original system has been refined and expanded [e.g., Coulthard & Montgomery, 1981], its central principles remain the same [Coulthard, 1985].)

Given the comprehensiveness of this ambitious system, it is not surprising that second language classroom researchers have tended to limit their analyses to specific aspects of discourse or to linear sequential descriptions of some of the acts present in classrooms (e.g., Fanselow, 1977), rather than attempting to specify which acts can realize particular moves, which sequences of moves compose different types of exchanges, how exchanges form a transaction, and how transactions fit together in a lesson. Given the fluidity of many classroom activities, such a grammar would indeed be difficult to write. A CALL program, on the other hand, structures the domain of possible discourse between student and computer, making it possible to write a grammar that describes the interaction.

DISCOURSE ANALYSIS OF CALL ACTIVITIES

Computer software makes possible two (though not mutually exclusive) discourse situations categorized on the basis of their participants. In the first instance, software can be used as the center of an activity that promotes conversation among human participants; the computer provides something interesting and dynamic to work on. Several researchers have examined the amount and
functions of language produced by learners in this situation (Piper, 1986; Mydlarski, 1987; Abraham & Liou, in press; Liou, 1989). Also examining communication between human participants, Esling (in press) proposes an investigation of the functional acts used in learners’ e-mail (i.e., electronic mail, the written messages people exchange via networked computers at remote locations).

This paper focuses on the discourse situation that occurs when the computer provides language practice through student-computer interaction. Depending on the program, the interaction allowed can render possible a variety of functional acts. A precise description of an activity could be formulated by specifying the types of acts possible within a given CALL program, which acts can be used as each type of move, how moves fit together to form legal exchanges, and so on, until the grammar of the CALL activity is defined. This grammar, then, provides an unambiguous statement of the parameters of student-computer interaction within a CALL program. The grammar of possible discourse forms a framework for describing actual acts of the students as they work, as well as a basis for comparison with the acts allowed in other CALL and classroom activities. To demonstrate how such an analysis is done, a CALL grammar lesson is first described anecdotally, then in the pedagogical and linguistic terms reviewed above, and finally using the principles from Sinclair and Coulthard’s discourse analysis system.

Description of a CALL Grammar Lesson

The CALL grammar lesson that serves as our example is one in a series of grammar and paragraph development lessons (Chapelle & Boysen, 1990) for intermediate- to advanced-level ESL students. The purpose of the lesson is for students to review and practice the correct forms and appropriate use of the present perfect tense in contrast to the past tense in a context requiring attention to both meaning and form. The main “page” of the lesson that the student sees on the computer screen in three colors (taken from the Phrases page of Grammar Lesson 6) is illustrated in Figure 4.

The page presents three groups of phrases: adverbial time phrases, subjects, and verb-plus-complement phrases. The student must choose a phrase from each group; as each phrase is chosen, the computer plots it on the lower portion of the screen, ultimately forming a sentence. However, the verbs are in the simple form. The student must edit the verb in each sentence, supplying the form appropriate to the given context. Because students require additional (factual) information to create meaningful sentences, help is available, labeled Facts and displayed in table format (see Figure 5). Ideally, the student will consult Facts to learn about the situation,
FIGURE 4  
Screen Display from Phrases Page Present Perfect and Past Tense

### TIME PHRASES

- Since the 1970s,  
  - In the 1980s,
  - During the prosperous 1980s,
  - Before the problems began in the 1970s,
  - After the problems began in the 1970s,
  - Since the rise in fuel prices,

### SUBJECTS

- car buyers
- the American car industry
- American cars

### VERBS AND COMPLEMENTS

- begin to want small cars.
- start to value fuel economy.
- purchase large cars.
- increase automation in the factory.
- have large engines.
- become fuel efficient.
- cut workers' wages.
- use all American auto parts.
- produce ears with V-8 engines.

Move the arrow and press RETURN to select a phrase from each section.  
Then, edit the verb to make it correct.

<table>
<thead>
<tr>
<th>PF1 Words</th>
<th>PF2 Facts</th>
<th>PF3 Grammar</th>
<th>PF4 Exit</th>
</tr>
</thead>
</table>

and will then be able to choose phrases that create factually true sentences, and to edit the verbs, rendering them correct and appropriate to the meaning of the sentences.

While all students are expected to need help with the facts of the lesson’s topic (the U.S. automobile industry), some students will also require definitions of certain words in the phrases, as well as a review of the verb forms and their uses. To receive help with words, students press PF1 (as indicated on the bottom of the screen); they are asked, “Which word?” They type the word, and the program returns its definition. To receive help with the verbs (PF3), students are offered two types of grammar help: They can look at several pages summarizing the grammar rules and forms for the present perfect and past tenses, or they can go through a step-by-step tutorial on how to form and edit the sentences in the exercise. All of the help—facts, words, and grammar—is optional; if students do not request to see these parts of the program, they never will. Instead, they can work on the exercise simply by forming and editing sentences, reading the computer’s evaluation of the sentences, and saving those they wish to print.

The computer evaluates and returns detailed feedback messages...
FACTS ABOUT THE PAST AND EVENTS THAT BEGAN IN THE PAST

<table>
<thead>
<tr>
<th>WHO/WHAT</th>
<th>DID WHAT?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WHEN? in the 1960s</td>
</tr>
<tr>
<td>Car buyers</td>
<td>Purchase large cars</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Car industry</td>
<td>Use American auto parts</td>
</tr>
<tr>
<td></td>
<td>Produce V-8 engines</td>
</tr>
<tr>
<td>American cars</td>
<td>Have large engines</td>
</tr>
</tbody>
</table>

(as illustrated in Figure 6) for every sentence that the student forms. When a sentence is incorrect, there are two possible problems: its meaning or its verb form. The meaning errors can be further subdivided into two types: those that create nonsense sentences, and those that create sentences that are not true according to the facts. Verb errors can also be subdivided into two: those in which the wrong tense or number is used given the subject and meaning of

FIGURE 6
Example Student Sentences and Computer Feedback

<table>
<thead>
<tr>
<th>Student’s sentence</th>
<th>Computer’s message</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Meaning errors</strong></td>
<td></td>
</tr>
<tr>
<td>Since the 1970s, car buyers have increased automation in the factory.</td>
<td>That doesn’t make sense because of the subject and verb. <em>Car buyers can’t increase automation in the factory.</em></td>
</tr>
<tr>
<td>During the prosperous 1980s, American cars became fuel efficient.</td>
<td>That’s not true. <em>During the prosperous 1980s</em> means PAST and that event continues today.</td>
</tr>
<tr>
<td><strong>Verb errors</strong></td>
<td></td>
</tr>
<tr>
<td>Before the problems began in the 1970s, car buyers have purchased large cars.</td>
<td>You must use the past tense (<em>purchased</em>) for this event in the past.</td>
</tr>
<tr>
<td>Since the 1970s, the American car industry has increased automation in the factories.</td>
<td>You must use the <em>increased form</em> after <em>has.</em></td>
</tr>
<tr>
<td><strong>Correct</strong></td>
<td></td>
</tr>
<tr>
<td>In the 1960s, American cars had large engines.</td>
<td>The verb is correct in that sentence.</td>
</tr>
</tbody>
</table>
the sentence, and those in which the verb is incorrectly formed according to the rules of English.

It is possible to describe this lesson using the parameters developed by Higgins (1988), Underwood (1984), and Phillips (1985). Figure 7 presents an interpretation of the parameters delineated under the magister/pedagogue and the communicative/noncommunicative distinctions.

FIGURE 7
The Grammar Lesson Described in Terms of

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Application to the grammar lesson</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pedagogical parameters</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Magister</strong></td>
<td></td>
</tr>
<tr>
<td>Program directs students' learning</td>
<td>No</td>
</tr>
<tr>
<td>Program has authority to evaluate, praise, censure</td>
<td>Yes</td>
</tr>
<tr>
<td>Program structures the order of events</td>
<td>No?</td>
</tr>
<tr>
<td>Program explains rules; gives examples</td>
<td>Yes</td>
</tr>
<tr>
<td>Program repeats endlessly</td>
<td>No</td>
</tr>
<tr>
<td><strong>Pedagogue</strong></td>
<td></td>
</tr>
<tr>
<td>Program assists students' learning</td>
<td>Yes</td>
</tr>
<tr>
<td>Program has no authority</td>
<td>No</td>
</tr>
<tr>
<td>Program provides no structure</td>
<td>No?</td>
</tr>
<tr>
<td>Program answers students' questions</td>
<td>Yes?</td>
</tr>
<tr>
<td>Program follows students' orders</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Linguistic parameters</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Noncommunicative CALL</strong></td>
<td></td>
</tr>
<tr>
<td>Program incorporates grammatical sequencing</td>
<td>No</td>
</tr>
<tr>
<td>Program judges to inform students of their errors</td>
<td>Yes</td>
</tr>
<tr>
<td>Program is in control</td>
<td>No</td>
</tr>
<tr>
<td>Subject matter is irrelevant to student</td>
<td>?</td>
</tr>
<tr>
<td>Lesson is predetermined</td>
<td>No?</td>
</tr>
<tr>
<td>Student perceives task as a required lesson</td>
<td>?</td>
</tr>
<tr>
<td>Student views task as identical to classroom activities</td>
<td>No?</td>
</tr>
<tr>
<td><strong>Communicative CALL</strong></td>
<td></td>
</tr>
<tr>
<td>Program does not impose grammatical sequencing</td>
<td>Yes</td>
</tr>
<tr>
<td>Program judges more to provide helpful hints</td>
<td>Yes</td>
</tr>
<tr>
<td>Student is in control</td>
<td>Yes</td>
</tr>
<tr>
<td>Student relates to subject matter in a personal way</td>
<td>?</td>
</tr>
<tr>
<td>Student creates own learning experience</td>
<td>Yes?</td>
</tr>
<tr>
<td>Student perceives task as motivating supplement</td>
<td>?</td>
</tr>
<tr>
<td>Student views task as a novel activity</td>
<td>Yes?</td>
</tr>
</tbody>
</table>

*Note:* ? indicates uncertainty of the value given or inability to assign a value because the parameter is too general or a matter of the student's opinion.
Overall, the lesson appears more pedagogue than magister and more communicative than not. The judgments one might make on the basis of Phillips’ descriptors are represented in Figure 8. All of these descriptions, however, are based on a single rater’s fallible judgment of the overall activity. In other words, these descriptors do not recommend themselves as criteria for a reliable, detailed view of a CALL activity.

**FIGURE 8**


<table>
<thead>
<tr>
<th>Parameter</th>
<th>Cloze exercise</th>
<th>Grammar lesson</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity type</td>
<td>Quiz</td>
<td>Exploratory?</td>
</tr>
<tr>
<td>Learning style</td>
<td>Recall</td>
<td>Recall, comprehension, experimental learning</td>
</tr>
<tr>
<td>Learner focus</td>
<td>Doing a test</td>
<td>Trying out sentences</td>
</tr>
<tr>
<td>Program focus</td>
<td>Control over syntactic form</td>
<td>Control over meaningful and grammatical sentences</td>
</tr>
<tr>
<td>Classroom management</td>
<td>Individual work</td>
<td>Individual work</td>
</tr>
<tr>
<td>Program difficulty</td>
<td>Little flexibility</td>
<td>Flexibility only in help options</td>
</tr>
<tr>
<td>Language difficulty</td>
<td>Could be a choice of levels</td>
<td>No flexibility, but help available</td>
</tr>
</tbody>
</table>

*Note.* ? indicates uncertainty of the value given or inability to assign a value because the parameter is too general or a matter of the student’s opinion.

**Discourse Analysis of the Grammar Lesson**

The problem of the unreliable and general character of these CALL descriptions may be solved by adapting the units of analysis and structures suggested by Sinclair and Coulthard for classroom discourse. At the lowest level of this ranked scale, the functional acts that the computer can perform are the following (listed in Figure 9): offer help (facts, words, grammar); offer a phrase; offer to let the student exit; require the student to edit; offer to save a sentence; exit from the transaction; save a sentence; provide help; add a phrase to a sentence; and judge a sentence. The student can perform the following acts: choose help, choose to exit, select a phrase, edit a phrase, and choose to save a sentence. The structure of these acts becomes apparent when each is assigned to a particular type of
move, which has a defined place in an exchange. According to the rules of this CALL activity, a move can be realized by only one act. The possible moves and acts that can realize them are detailed in Figure 9. As an example, the initiating move should be read as follows: The computer can perform an initiating move that can be realized as offering help; offering to let the student exit; offering a phrase; requiring the student to edit; or offering to save a sentence.

FIGURE 9
The Grammar Lesson Participants, Moves, and Acts

<table>
<thead>
<tr>
<th>Participant</th>
<th>Move</th>
<th>Acts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer</td>
<td>Initiate</td>
<td>Offer help, offer exit,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>offer phrase, require edit,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>offer save sentence</td>
</tr>
<tr>
<td>Computer</td>
<td>Obey</td>
<td>Exit, save sentence</td>
</tr>
<tr>
<td>Computer</td>
<td>Follow Up</td>
<td>Add phrase, judge, provide help</td>
</tr>
<tr>
<td>Student</td>
<td>Choose</td>
<td>Choose help, choose exit,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>choose save sentence</td>
</tr>
<tr>
<td>Student</td>
<td>Respond</td>
<td>Select phrase, edit</td>
</tr>
</tbody>
</table>

The structure of interaction is defined by the grammar of this discourse analysis system: In this lesson, these moves fit together in triplets called exchanges. There are three types of exchanges (see Figure 10) allowed in this lesson: teaching, focusing, and framing. A teaching exchange begins with the computer initiating, or asking for, a response. Next, the student responds—either selecting a phrase, or editing a phrase. Then, the computer follows up—either adding a phrase to the sentence, or judging the student’s sentence. All of the acts in the teaching exchange are directly related to the instructional task at hand, rather than to the management of the program or to gaining additional information. A focusing exchange, on the other hand, is used to provide students with additional help. It also begins with an initiating move (made by the computer). The initiating move is followed by the student’s choice for help (facts, words, or grammar). The follow-up move is realized by the computer providing the corresponding help. A framing exchange is used for program management: saving sentences and ending a transaction. It, like the others, begins with the computer’s initiating move, which offers an option. The choice made by the student is to
save the sentence or to exit. The computer obeys by saving the sentence or exiting from the transaction, then beginning the next one.

FIGURE 10
The Grammar Lesson Exchanges and Moves

<table>
<thead>
<tr>
<th>Exchange</th>
<th>Moves (in the order indicated)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teaching</td>
<td>Initiate, respond, follow up</td>
</tr>
<tr>
<td>Focusing</td>
<td>Initiate, choose, follow up</td>
</tr>
<tr>
<td>Framing</td>
<td>Initiate, choose, obey</td>
</tr>
</tbody>
</table>

In Sinclair and Coulthard’s classroom description, the teaching and framing exchanges are similar to the ones described for the grammar lesson. The teaching exchange is used for teachers’ questions, students’ responses, and teachers’ evaluations. The framing exchange is usually used to mark the boundaries of transactions. However, the focus exchange is usually “a metastatement about the transaction” (Coulthard, 1985, p. 123), a definition that necessitates the instructor’s control over the transaction. The focus exchange in the grammar lesson, in contrast, is under the control of the student, who will choose to focus on one or more aspects of help during a transaction.

In Sinclair and Coulthard’s system, a transaction is defined as a sequence of exchanges. Transactions within the student-computer interaction can be defined in terms of particular obligatory and optional exchanges, as indicated by this example that describes the grammar lesson:

(FOCUS)* TEACH (FOCUS)* TEACH (FOCUS)* TEACH TEACH ((FRAME)* (TEACH)*)* FRAME

The parentheses indicate optional exchanges; the asterisk denotes an unlimited number of an exchange type. The grammar of this example reads as follows: A transaction consists of an optional, unlimited number of focusing exchanges, followed by an obligatory teaching exchange, followed by an optional, unlimited sequence of focusing exchanges, followed by a teaching exchange, followed by an optional, unlimited number of focusing exchanges, followed by two obligatory teaching exchanges, followed by any of these three possibilities: (a) an optional, unlimited number of framing exchanges; (b) an optional, unlimited number of teaching exchanges; or (c) an optional, unlimited number of framing and...
teaching pairs. The transaction must be terminated by a framing exchange. These transactions, unordered and of unlimited number, form a lesson, noted as follows:

**TRANSACTION (TRANSACTION) **

This discourse analysis system defines the specific acts that can occur and the structure of discourse in which they can occur in CALL activities. In doing so, it provides a level of detail that is much more precise than terms such as learner-controlled and exploratory. These two terms can be more precisely defined by specifying (a) the number and functions of the acts performed by the student and computer (in the case of the grammar lesson as defined in Figure 9, there are 5 for students, and 10 for the computer); (b) the functions of the computer’s follow-up moves (add phrases, judge sentences, and give help); (c) the ratio of optional-to-obligatory exchanges in a transaction (1-to-1); (d) the number of obligatory and optional transactions in a lesson (1 obligatory and unlimited optional); and (e) the order of the transactions in a lesson (no order). (Note that the simple notation introduced here is not adequate for expressing ordered transactions.) Despite the objective detail of this analysis, for reporting research on students’ CALL use, it alone has little validity because it describes only what students can do while they work on the grammar lesson; it fails to report what students actually did.

**Discourse Analysis of Student Data**

An empirically based description can be constructed using the structure of possible discourse as a framework for plotting what students actually did while they were working on the lesson. Of course, because some exchanges are obligatory, their acts must be performed by any student who completes the activity; however, many exchanges are optional, as is the number of transactions in this lesson. Consequently, in reporting research based on this activity, essential to the description of the materials will be the quantity and quality of acts that subjects actually performed. The efficacy of this approach is illustrated by data from three subjects in a pilot group who worked on this grammar lesson. This information was obtained from each student’s on-line file, which automatically stores information about what students are doing while they are working on the CALL materials. A protocol of Student A’s interaction is presented (in Figure 11) to demonstrate the kind of discourse that took place. The student began by choosing phrases composing the sentence, “Since the 1970s, car buyers begin to want new cars.” This
FIGURE 11
Protocol Data from Student A: Choices, Editing, and Feedback

<table>
<thead>
<tr>
<th>Participant</th>
<th>Action</th>
<th>Choices, editing, feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student</td>
<td>Choose</td>
<td>Since the 1970s, car buyers begin to want new cars.</td>
</tr>
<tr>
<td>Student</td>
<td>Edit</td>
<td>begin → began</td>
</tr>
<tr>
<td>Computer</td>
<td>Judge</td>
<td>You need the present perfect (with have) for this event because it happened in the past and continues now.</td>
</tr>
<tr>
<td>Student</td>
<td>Edit again</td>
<td>began → have begun</td>
</tr>
<tr>
<td>Computer</td>
<td>Judge again</td>
<td>Use the begun form after have</td>
</tr>
<tr>
<td>Student</td>
<td>Edit again</td>
<td>have began → have begun</td>
</tr>
<tr>
<td>Computer</td>
<td>Judge again</td>
<td>The verb is correct in that sentence.</td>
</tr>
<tr>
<td>Student</td>
<td>Choose</td>
<td>During the prosperous 1960s, car buyers purchase large cars.</td>
</tr>
<tr>
<td>Student</td>
<td>Edit</td>
<td>purchase → purchased</td>
</tr>
<tr>
<td>Computer</td>
<td>Judge</td>
<td>The verb is correct in that sentence.</td>
</tr>
<tr>
<td>Student</td>
<td>Choose</td>
<td>For the past several years, car buyers begin to want small cars.</td>
</tr>
<tr>
<td>Student</td>
<td>Edit</td>
<td>begin → began</td>
</tr>
<tr>
<td>Computer</td>
<td>Judge</td>
<td>You need the present perfect (with have) for this event because it happened in the past and continues now.</td>
</tr>
<tr>
<td>Student</td>
<td>Edit again</td>
<td>began → have begun</td>
</tr>
<tr>
<td>Computer</td>
<td>Judge</td>
<td>The verb is correct in that sentence.</td>
</tr>
<tr>
<td>Student</td>
<td>Choose</td>
<td>In the 1960s, car buyers begin to want small cars.</td>
</tr>
<tr>
<td>Student</td>
<td>Edit</td>
<td>begin → begin [Student chose not to edit.]</td>
</tr>
<tr>
<td>Computer</td>
<td>Judge</td>
<td>That's not true. In the 1960s means past and that event continues today.</td>
</tr>
<tr>
<td>Student</td>
<td>Edit</td>
<td>begin → began</td>
</tr>
<tr>
<td>Computer</td>
<td>Judge</td>
<td>That's not true. In the 1960s means past and that event continues today.</td>
</tr>
</tbody>
</table>

(Student tries two more verb forms, receives the same message each time, then exits.)
student then edited the verb as shown; he changed it to “began.” The verb tense was incorrect, given the meaning of the sentence; the computer therefore responded: “You need the present perfect (with have) for this event because it happened in the past and continues now.” The student reedited the verb twice, then continued, as indicated in Figure 11.

The data obtained from Student A’s work on the lesson, as well as the data from two other students (Figure 12) illustrates how their work can be concisely expressed. Not surprisingly, it is apparent that the three students took different approaches to the task. Only Student B selected a help option. Student A quit after creating a

**FIGURE 12**
Example Data from Students A, B, and C

<table>
<thead>
<tr>
<th>EXCHANGES</th>
<th>MOVES</th>
</tr>
</thead>
<tbody>
<tr>
<td>(FOCUS)*</td>
<td>+</td>
</tr>
<tr>
<td>TEACH</td>
<td>C</td>
</tr>
<tr>
<td>FRAME</td>
<td>G</td>
</tr>
</tbody>
</table>

*EXCHANGES*
(FOCUS)* = Optional, unlimited focusing exchange composed of initiating, choosing, and follow-up moves
TEACH = Obligatory teaching exchange composed of initiating, responding, and follow-up moves
FRAME = Obligatory framing exchange composed of initiating, choosing, and obeying moves

*MOVES*
+ = Neutral move made by the student or computer (neither correct nor incorrect, nor was it judged so by the computer)
C = Correct response or follow-up indicating correctness
X = Grammatical error in a response or a follow-up indicating a grammatical error
M = Meaning error in a response or a follow-up indicating a meaning error
G = Choice of grammar help

**Student A**
(FOCUS) *TEACH (FOCUS) *TEACH (FOCUS) *TEACH TEACH (TEACH)(TEACH)(TEACH) *FRAME +++ +C+ +XX +XX +CC +++
(FOCUS) *TEACH (FOCUS) *TEACH (FOCUS) *TEACH TEACH (TEACH)(TEACH)(TEACH) *FRAME +++ +C+ +CC +++
(FOCUS) *TEACH (FOCUS) *TEACH (FOCUS) *TEACH TEACH (TEACH)(TEACH)(TEACH) *FRAME +++ +C+ +XX +CC +++
(FOCUS) *TEACH (FOCUS) *TEACH (FOCUS) *TEACH TEACH (TEACH)(TEACH)(TEACH) *FRAME +++ +M+ ++M ++M ++M ++M +++

Note. Parentheses indicate optional exchanges; asterisks denote an unlimited number of exchange types.
Corresponds to protocol data in Figure 11.
sentence whose meaning was incorrect without trying to correct it or form another sentence. Student C, on the other hand, did not quit until having a sentence judged correct. The data also exhibit some similarities. None of the students chose help immediately after making errors. Student B used help once, but then exited immediately thereafter. They all appear to have been attempting correct answers with each of their responding moves; whenever they received a grammar error message from the computer, they kept working on the sentence until they got it correct.

This summary of data permits an accurate, detailed description of how students work with the lesson. More detail could be added to this summary of data: One could, for example, indicate which phrases were chosen by numbering each phrase and recording those chosen in the students’ profiles. Even with the small amount of data...
illustrated here, we can now begin to address the question: Do these data support the general descriptions of the lesson such as those developed using the parameters of Higgins (1988) and Phillips (1985)? For example, was it accurate to label this primarily a learner-controlled (pedagogue), exploratory activity? From the limited data examined, the three learners did use some different acts, to form somewhat different structures of interaction—structures that they themselves controlled. Yet, the number of transactions is similar for all three subjects. What does learner-controlled mean if all learners choose to use the software in a similar fashion? Grammatically speaking, these three subjects did not appear to explore the verb tenses. After a single instance of creating a correct verb form, they exited from the transaction and began another. This is contrary to what one would expect of students who were using the exercise as an exploratory activity. This analysis suggests that it may not be appropriate to term software exploratory; rather, it seems that exploratory as a term must refer to the behavior of the students.

COMPARING CALL ACTIVITIES TO OTHER INSTRUCTIONAL ACTIVITIES: A FUTURE RESEARCH AGENDA

In looking at the actual data from the “exploratory” program, one might wonder if they define the limits of student exploration with CALL. To answer this question, it would be necessary to examine data from a larger number of subjects working with this software, as well as to compare the actual data from the subjects who worked on this activity with data obtained from students working on other CALL activities. The concepts adapted from Sinclair and Coulthard, along with the notation introduced here, comprise the foundation of the formal system needed to make such comparisons. It is a formal system because it has a syntax (e.g., exchanges are composed of moves; moves are composed of acts) and a semantics (acts are interpreted as particular, real-world phenomena) of its own. Within the rules of the system, elements can be expanded and rearranged to express the variety of student-computer interactions that may take place. Expression of different activities using the same formalism is a necessary first step in making comparisons, as noted by discourse analysts (Coulthard, Montgomery, & Brazil, 1981) using Sinclair and Coulthard’s principles:

We develop new descriptions for each situation using the same descriptive principles. Thus we were engaged in describing what it is that makes TV discussions, for example, different as speech events from
doctor/patient interviews and committee meetings, but of course for a set of such descriptions to be mutually enlightening they would need to have more or less common modes of description. (p. 15)

That common mode of description is also what is needed for expressing student-computer interaction.

In comparing one computer activity to another, then, it is necessary to identify the functional acts the activity allows, and the structure of those acts. For example, in the cloze exercise summarized using Phillips’ (1985) descriptors in Figure 3, we might use this discourse analysis system to delineate what functional acts are allowed, which moves they realize, how those moves form exchanges, and what the optional and obligatory exchanges are in a transaction. If the cloze activity were expressed using the same formal system as the grammar lesson, their possible discourses could be compared, and researchers reporting results based on the two activities could use one set of expressions to summarize actual acts performed by subjects using both activities. Of course, more acts will need to be introduced to describe other CALL activities.

This paper is intended to illustrate these principles rather than to propose the set of necessary acts to account for CALL activities. However, one might ask if it is necessary to use the complex system devised for human interaction to express and compare the limited discourse of CALL activities. Indeed, current pedagogical and linguistic descriptors for CALL are presented as unique from other language classroom activities. However, to isolate the constraints a CALL activity places on discourse in contrast to other language learning activities, it is necessary to use terminology comparable to what may be used for classroom discourse. For example, we note that the moves a student could perform in the grammar lesson were choosing and responding (each realized as the acts listed in Figure 9). We will require a shared formalism to answer the following types of questions: How do the quality and quantity of these moves (and the acts that can realize them) compare to those a student can perform in various classroom activities? How do the types of judging acts (feedback messages about meaning and grammar) used as the computer’s follow-up moves compare to the acts Chaudron (1977) observed language teachers using as follow-up moves in classes? How do the task-defined variations on the initiation-response-feedback classroom exchange described by Heap (1988) compare to their counterparts in a CALL activity?

Unfortunately, as noted above, the structure of classroom discourse in language classroom research is typically not as clearly stated as in Chaudron’s (1977) analysis. However, attempts to synthesize this research (Chaudron, 1988) have clarified the problem of
disparity among research methods, thereby encouraging future choices of discourse categories with an eye toward linguistic comparisons. A future agenda for language classroom research may work toward a standard formalism for describing and comparing activities with one another in terms comparable to Sinclair and Coulthard’s. Until that time, however, the generalizability of CALL research results can be clarified by using this standard formalism to describe the specifics of students’ actual interaction with the computer. Such a discussion accompanying results of learning outcomes or attitudes of particular students will offer the detail necessary for assessing the relevance of research results in one study to the expectations for CALL activities in another situation.

CONCLUSION

The primary focus here has been the analysis of discourse units and their structures in CALL activities in order to improve the internal and external validity of CALL research, thereby strengthening our understanding and, ultimately, the classroom use of CALL. In fact, second language researchers as well as CALL theorists and developers will find such an analysis useful for other purposes.

Second language researchers attempting to assess the effects of a teaching method that remains constant for all students might offer students instruction using a formally described CALL program. The need for such consistency is exemplified by descriptive studies (e.g., Chaudron, 1977; Allwright, 1975) documenting the inconsistency of teachers’ treatment of errors. This spontaneity, although obviously an asset in many ways, renders difficult the controlled study of the relative benefits of specific aspects of instruction (acts or moves) of theoretical interest. A computer program, on the other hand, maintains the necessary consistency. Robinson and Others [sic] (1986) exploited this capability, using a CALL program to test, among other hypotheses, Krashen’s (1982) theory concerning implicit error correction. However, without a precise and definable formalism, the report of results fails to detail the actual student-computer interaction; moreover, missing from such research is a clear specification of the relationship between media-supported instructional treatments and their regular classroom counterparts. With an overall discourse framework using terminology common to both contexts, it may be possible to reinvestigate the use of the computer for such treatment experiments.

Additionally, second language researchers investigating students’ language learning strategies may be interested in the acts that
students perform while they are working on various activities. Individual or combinations of acts can be used as reliable definitions of language learning strategies. An obvious strategy used by some students working on the grammar lesson when they opted for a choose help act was resourcing—the use of target language reference materials (as defined by O’Malley, Chamot, Stewner-Manzanares, Kupper, & Russo, 1985). Other strategies that may be defined through the use of combinations of acts in the grammar lesson are planning (indicated by relationships among choose phrase acts), monitoring (displayed by the consistent use of editing acts rather than reliance on the computer’s judging acts for corrections), and practice (evidenced by the number of transactions completed). Some strategies have been assessed successfully through the use of computer-collected data from CALL programs (Jamieson & Chapelle, 1987; Mizuno, 1989); however, such results will be more comprehensible and generalizable if definitions of investigated strategies are phrased in terms of students’ acts performed within a defined domain of possible discourse.

Examination of variation in students’ acts within the framework of defined contexts (types of exchanges, transactions, and lessons) holds a key to long-awaited progress in research on the relationship between types of instruction and particular student characteristics. Skehan (1989) notes that surprisingly little work has been done in this essential element of a theory of second language acquisition. Indeed, CALL lessons have been used as a means for presenting material to students in a defined manner to observe outcomes for different types of learners (e.g., Abraham, 1985); however, to interpret and generalize results, so that future research can build on current findings, it is necessary to express instructional treatments in precise, empirically based, formal terms.

CALL theorists and developers are also in need of precision and consistency in their terminology. Discussion of types of CALL can proceed with maximum clarity if specific terms for CALL interaction are used. The general terms learner-controlled and exploratory for the present perfect lesson were restated more precisely and objectively in terms of discourse units and structures. Discussion of other general terms of interest such as intelligent programs would be clarified by an objective definition as well—a definition based on the number and quality of the acts that realize the computer’s follow-up moves, perhaps. Such a definition would provide a concrete basis for exploring development of intelligent programs and for observing learners’ use of aspects of a program’s “intelligence.”

Use of the same principles for student-computer interaction as for
human interaction applies insights from linguistic research, thereby opening a novel perspective on unsolved human-factors issues in CALL. Anyone who has witnessed first-time computer users notes the halting flow of the interaction as they turn to the instructor to ask, “What should I do now?” rather than reading the instructions on the screen. Culley, Mulford, and Milbury-Steen (1986) cite protocol data from students working on an activity similar to a CALL adventure game. The developers intended to have students enter commands to the computer, telling the program to take them to the front hall, open the letter, etc. Students had trouble understanding their role in the dialogue. These and other observations can be viewed more clearly from the perspective of linguists who note elements of human conversations that facilitate smooth exchanges (Sacks, Schegloff, & Jefferson, 1974). In light of such factors, computer software can be assessed and improved in terms of its performance as a participant in a conversation.

To maximize the utility of the CALL discourse analysis presented here, it will ultimately be necessary to add several dimensions. First, for a clear perspective on the quality of functional acts allowed, acts should be categorized in terms such as Halliday’s (1977) ideational, interpersonal, and textual functions; Brown and Yule’s (1983) transactional versus interfactional functions; and other terms that may be significant for second language acquisition research. A second dimension to be added is greater specificity of the acts and their sequencing. For example, in the grammar lesson, there are particular combinations of choose phrase and edit verb acts that precede particular judge acts. An important part of the empirical description will be to note which of those choosing, editing, and judging acts are actually used by students. Finally, the analysis of the units and their structure leave much room for addition of contextual-pragmatic considerations (cf. Halliday & Hasan, 1989). The functions and structure presented here do not represent a comprehensive system, equipped to analyze all relevant aspects of CALL activities; instead, they provide a first step toward focusing a research agenda that promises to support progress in second language acquisition research and research in CALL.
ACKNOWLEDGMENTS

This project was supported in part by the research program in the Department of English at Iowa State University. I am grateful to Pete Boysen and Rex Thomas for their essential support and assistance.

THE AUTHOR

Carol Chapelle is Associate Professor of English at Iowa State University, where she teaches courses in TESL and applied linguistics. She conducts research on computer-assisted language learning and second language acquisition.

REFERENCES


Liou, H-c. (1989, March). Native language and target language functions of pairs of students working with computer software. In C. Chapelle (Chair), Descriptive research approaches for CALL activities. Colloquium conducted at the 23rd Annual TESOL Convention, San Antonio, TX.


AS AN INTERNATIONAL FAMILY

25 YEARS

1991 TESOL Convention

at the
New York Hilton
and
Sheraton Centre
New York City
New York, USA
March 24-28, 1991

- Plenaries
- Papers
- Workshops
- Colloquia
- Educational Visits
- Job Placement

Teachers of English
to Speakers of Other Languages, Inc.

an international professional association for teachers of English as a second or foreign language, and bilingual education, as well as for researchers of language acquisition, language theory, and language teaching pedagogy.

Program Chair: Suzanne Griffin, Office of the State Superintendent of Public Instruction, Olympia, Washington, USA

Associate Program Chair: Mary Ann Christison, Snow College, Ephraim Utah, USA

Members of TESOL will receive convention information. Others may contact:

TESOL Convention Department
1600 Cameron Street, Suite 300 Alexandria, Virginia 22314-2705 USA
Telephone 703-836-0774 Fax 703-836-7864
Predicting Success for 
International Teaching Assistants 
in a U.S. University

GEORGE YULE and PAUL HOFFMAN
Louisiana State University

In this analysis of the performance of 233 international graduate assistants during a 2-year period, we attempted, via Test of English as a Foreign Language (TOEFL) and Graduate Record Exam (GRE) scores submitted at the time of application, to predict which of these students would eventually receive positive or negative recommendations to be assigned teaching duties. Students who received negative recommendations were found, on average, to have significantly lower TOEFL and GRE Verbal scores than those who received positive recommendations. The percentage of each recommendation group scoring at or above a series of TOEFL cutoff scores was established and used to calculate the ratio of risk (funding students who will receive a negative recommendation) to reward (funding students who will receive a positive recommendation). The relationship of recommendation type to subsequent grade point average (GPA) showed a significant difference in favor of the positive group during the first year of graduate study, but not thereafter. Implications are explored for decision making and the advising of other academic departments regarding the awarding of teaching assistantships to international students.

The increasing number of nonnative-speaking international teaching assistants (ITAs) in U.S. universities has resulted in an unprecedented demand for training/preparation programs from ESL educators. Their advice actively sought by university vice-chancellors and deans, their programs mandated by state legislatures, and their opinions reported in the national press (cf. contributions in Bailey, Pialorsi, & Zukowski/Faust, 1984, and Chism, 1987), the ESL professionals involved in such programs have found themselves in a political spotlight unusual for the field. One effect of this attention has been a tendency to talk of solutions
to the perceived problem, widely known as “the foreign TA problem” (Bailey, 1984), in programmatic terms. Reports by the professionals involved tend to describe a particular program, discussing how it was established, how long it lasts, and what kinds of activities are used. Because the ESL professional (teacher or administrator) is in the spotlight, the most typical report focuses on the activities of the ESL professional (cf. Byrd & Constantinides, 1988; Constantinides, 1987; Constantino, 1985; Kinney & Brown, 1988; Rice, 1984; Schneider & Stevens, 1987; Sequeira & Darling, 1987). Occasionally, and more frequently of late, there have been studies reporting on the international teaching assistant as an ESL speaker in an instructional setting (cf. Ard, 1986; Douglas & Selinker, 1987; Rounds, 1987a, 1987b; Tyler, 1988; vom Saal, 1988).

However, it is much less common to find a report that focuses on the basic administrative issue of awarding teaching assistantships on a responsible and reliable basis. The TESL literature has generally described what has been done to solve a problem created by other academic disciplines; it has not typically been in a position to mandate specific proficiency levels for assistantship awards. Indeed, in many cases, the ESL program sustains full responsibility for improving the spoken instructional performance of those ITAs already selected by departments, but very little power to determine who must take an ITA training course and who may or may not be assigned teaching duties. (Note that in Landa & Perry’s [1984] study, some ITAs with negative evaluations were themselves allowed to make the decision to continue teaching.) In some situations, however, the ESL program has been granted the authority to determine policy regarding ITA training and assignment of teaching duties. This context provides ESL professionals the opportunity to develop guidelines for other academic disciplines concerning the likely future success of an ITA on the basis of application data. If our expertise as second language educators is sought to solve a general ITA problem, we should be able to quote patterns of likely success or failure based, to a certain extent, on the international student’s English language skills before an assistantship is awarded. The study reported here was conducted to establish such patterns.

BACKGROUND

Like most large public universities, Louisiana State University provides a preparation course for ITAs. The course could be described as “seminar-type” rather than “orientation-type” (Turitz, 1984, p. 43). More specifically, following the topological
distinctions of Constantinides (1987), it would qualify as “a term-
long course prior to teaching” (p. 277), mandated by the Graduate
School for all newly arrived international graduate students who
have been awarded assistantships. At Louisiana State University,
the assistantship awards are made prior to the student’s arrival in the
U.S. on the basis of their application materials, and are not tied to
specific duties such as research, service, or teaching. The students
receive their assistantship stipends while taking the course, but are
not permitted to undertake any teaching duties until recommended
to do so at the end of the course.

The three-credit-hour course lasts 15 weeks and incorporates
many of the features that have become typical of ITA preparation
courses, such as exercises on segmental and suprasegmental aspects
of English; discussion of presentation skills and aspects of cross-
cultural communication; a range of one-way (information transfer)
and two-way (interactive) tasks; and a number of audiotaped and
videotaped performances under different conditions, with both
group and individual feedback sessions (cf. Byrd, Constantinides, &
Pennington, 1989, for typical materials). At the end of the course, a
recommendation (positive or negative) is made regarding each
student’s ability in spoken English, with particular reference to the
oral presentation of instructional material. A positive recommenda-
tion permits the student to be assigned teaching duties. Those
receiving a negative recommendation cannot be assigned any
teaching duties.

This recommendation (plus the responsibility for defending its
validity) is made by a faculty member in the Linguistics Program on
the basis of a profile of each student developed over the semester.
That profile is determined by the results of a form of “performance
testing” (Bailey, 1984, p. 13), which involves (a) viewing videotapes
of each student in an instructional role; (b) gauging the opinions of
each student’s ESL instructor, particularly regarding performance
in tasks not videotaped; and (c) noting the results of undergraduate
student evaluations. (The undergraduates make these evaluations as
members of a group receiving instruction from the ITA. See Yule &
Hoffman, in press, for more details.)

Having implemented and administered this course for 2 years,
we believed that we had substantially improved the situation of
ITAs on campus; on the one hand, because of the semester-long
acculturation that accompanied the course, and on the other,
because of the substantial reduction (via negative recommenda-
tions) in undergraduates’ experiences of ITAs whose English was
extremely difficult to understand. However, we were concerned
that so many assistantships were being awarded to students who
continued to have substantial difficulty in spoken English, even after a semester of instruction in our program and some immersion in the life of a U.S. university. Perhaps there had been clues in the application data that might have been used more carefully to screen those applicants to whom assistantships, particularly assistantships that would involve teaching duties, were being awarded. To find those clues, we attempted to profile retrospectively the “at time of application” characteristics of those students who went on to receive a positive recommendation versus those who later received a negative recommendation.

THE STUDY
Unlike Abraham and Plakans (1988), who reported on the screening power of the SPEAK test, we had no comparable objective measure of the students’ oral proficiency prior to their arrival. In any case, since the SPEAK test is administered after the students have already arrived at their universities, those scores would not be available at the time of the student’s application. Like many other U.S. institutions, we had had very little success in our earlier attempts to require the Test of Spoken English (TSE) scores from international graduate students as part of their application for assistantships, perhaps because of the costs involved (cf. Mellor, 1987), or simply because “this test continues to be inaccessible for many international students around the world” (Schneider & Stevens, 1987, p. 286). What we did have, in common with every graduate school in the U. S., were TOEFL and Graduate Record Exam (GRE) scores. Since it is the combination of the GRE Verbal component (GREV) and the GRE Quantitative component (GREQ) that graduate admission committees typically use in setting minimum GRE levels for nonprobationary admission, we elected to use those two component scores separately, along with the TOEFL scores, as possible predictors in our study.

RESULTS AND DISCUSSION
Using TOEFL and GRE Scores
Table 1 provides summary statistics (from application data) for the whole population and also for those two subgroups who later (after a semester-long course) received positive or negative recommendations. For the whole group, a strong correlation ($r= .76, p < .001$) was found between TOEFL and GREV, with weaker correlations between TOEFL and GREQ ($r = .13$) and between GREV and GREQ ($r= .27$).
TABLE 1
Summary Statistics for the Whole Group and Two Different Recommendation Subgroups

<table>
<thead>
<tr>
<th></th>
<th>Whole group (N = 233)</th>
<th>Recommendation subgroup</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>TOEFL</td>
<td>593</td>
<td>42</td>
</tr>
<tr>
<td>GREV</td>
<td>447</td>
<td>108</td>
</tr>
<tr>
<td>GREQ</td>
<td>712</td>
<td>66</td>
</tr>
</tbody>
</table>

When we look at the two subgroups, there is some indication that, on those measures that one would associate with English language proficiency (i.e., TOEFL and GREV), these two groups were certainly distinguishable, but that there was no distinction in terms of GREQ. The difference between mean TOEFL scores for the positive and negative recommendation groups was statistically significant ($t_{[231]} = 9.34, p < .001$), as was the difference between mean GREV scores for the two groups ($t_{[231]} = 6.44, p < .001$). The minor difference noted between mean GREQ scores for the two groups did not reach statistical significance ($t_{[231]} = .94, n.s.$).  

Recommendation status was predicted from entrance TOEFL and GREV scores using multiple logistic regression (Cox, 1970) estimated by maximum likelihood. TOEFL had a significant effect (the regression coefficient was .0358) on the recommendation status ($\chi^2[1, N = 233] = 25.9, p < .0001$). That is, each one-point increase in a TOEFL score corresponds to a 3.6% increase in the odds of being recommended to teach. There was no additional significant effect of GREV (the regression coefficient was .018) on recommendation status ($\chi^2[1, N = 233] = .41, p < .52$). Clearly, the TOEFL scores at application do have some predictive power with regard to eventual recommendation status for our population.

Although one might not anticipate an extremely strong relationship between performance on a predominantly written language test such as the TOEFL and a later performance judged in terms of the clear and comprehensible presentation of instructional material in spoken English (cf. Clark & Swinton, 1979), this should not lead us to conclude that a TOEFL score can provide no basis for judging the level of future spoken-English performance. Our data suggest

---

1 Although multiple t-tests pose a recognized problem, in this case they do not: Only three t-tests are involved and the significant p-values are so small.
that there is something valuable to be learned from the TOEFL’S predictive power with regard to our population. From a practical point of view, for example, we might want to have some idea of relationships existing between specific TOEFL scores and eventual positive or negative recommendations. That is, we might benefit from knowing what likelihood of receiving a positive or negative recommendation appears to be associated with specific TOEFL scores. This information can be presented in different ways.

Given the similar pattern in group differences, as shown in Table 1, and the strong correlation between TOEFL and GREV scores, which suggest that both these measures may provide some guidance in trying to anticipate eventual recommendation, we shall illustrate the patterns for both test scores in Figures 1 and 2. While the GREV patterns may prove to be useful in the absence of a TOEFL score (or in those situations where the TOEFL score is sufficiently old and therefore suspected of not being a reflection of current proficiency),

FIGURE 1
Percentage of Each Recommendation Group At or Above a Series of TOEFL Scores

Recommendation
'Negative
• Positive
we shall concentrate our discussion on the patterns illustrated by the more commonly used TOEFL measures.

Figure 1 represents the distribution of the students in the negative and positive recommendation groups, according to the percentage of students in each group who had scored at or above the specific TOEFL scores indicated. Figure 2 presents comparable data for GREV scores. Since most institutions would generally wish to avoid awarding assistantships to students who go on to receive negative recommendations, we can take the information represented in Figures 1 and 2 and display it as a regression line (Figures 3 and 4), allowing us to see what percentage of the whole population received a negative recommendation at a series of TOEFL and GREV cutoff scores.

In Figure 3, we have a graphic answer to the question: What percentage of this population of international graduate students, identified via TOEFL score upon application, eventually received a negative recommendation? Using a TOEFL score of 530 (and above) as a prerequisite for an assistantship award, we would have

FIGURE 2
Percentage of Each Recommendation Group
At or Above a Series of GREV Scores
included all of those who received a negative recommendation (i.e., about 30% of the population in this study). Using a TOEFL score of 570 (or higher) as a prerequisite, we would have reduced by half the number of those receiving negative recommendations. The clear pattern is that the percentage of students receiving a negative recommendation decreases fairly dramatically as the TOEFL score increases. It would appear, from these data, that one way of avoiding negative recommendations completely would simply be to make higher TOEFL scores, such as 625 (and above), a requirement for assistantship awards.

While this apparent solution to the negative recommendation issue would seem to offer program administrators a means of minimizing their risks when awarding assistantships, it fails to consider the results of our study from a different perspective, one that reveals the existence of another kind of risk. If the TOEFL cutoff score had been set fairly high for the population in our study, then the outcome would have been that many students who
eventually went on to receive positive recommendations would have been excluded from assistantships. In Figures 5 and 6, we represent the percentages of each population that would have been excluded from assistantship awards if different TOEFL or GREV cutoff scores had been used. (Note that this is the reverse of the perspective taken in Figures 1 and 2.) Essentially, we are trying to answer the administrator’s question: “If I award an assistantship to an international graduate student applicant with a TOEFL score $x$, what is the likelihood that I will avoid giving the assistantship to someone who would go on to receive a negative recommendation (i.e., to my benefit), while running the risk of failing to award an assistantship to someone who would go on to get a positive recommendation (i.e., to my detriment)?”

Figure 5 illustrates that, as we increase our TOEFL cutoff score to exclude a larger number of those who would receive negative recommendations, we concurrently exclude an increasing number
of those who would receive positive recommendations. From Figure 5, we can calculate that if we set a TOEFL score of 550 as our cutoff level for assistantship awards, we will exclude about 40% of those who would go on to receive negative recommendations and also exclude about 7% of the positive recommendation group. At a 570 TOEFL cutoff level, we will exclude about 70% of the negative group and just over 20% of the positive group. Notice that, at around 630 on the TOEFL scale, we would exclude all of the negative group, and retain almost 40% of the positive group.

Another way to represent the total “cost” of making a wrong decision at application time is via a scatterplot (Figure 7) showing the percentage accepted of the negative group relative to the percentage rejected of the positive group at a series of TOEFL cutoff scores. As can be deduced from Figure 7, there are cutoff points that would have to be used if it were more than twice as risky to accept a negative as to reject a positive (i.e., around 600). However, if it is, in terms of acceptable risks, less than twice as bad
to accept a negative as to reject a positive, then the cutoff can be set closer to 560.

The obvious implication of the ratios in Figures 5 and 7 is that, if an academic program has funding only for assistantships that necessarily involve teaching duties, then avoiding negative recommendations would be extremely important and higher entry TOEFL scores would have to be sought. On the other hand, if a program has experienced some difficulty in recruitment, yet has funding for assistantships that are not specifically tied to teaching duties (and could therefore absorb the impact of some negative recommendations), then lower entry TOEFL scores could be tolerated. Consequently, the optimum ratio of risk-to-reward must be determined by each academic program. The patterns represented in the figures shown here provide, at the very least, a potentially informative basis for making that determination.

Using Grade Point Averages

The concept of “success” for an international graduate student
attending a U.S. university is not only a matter of holding a teaching assistantship; it is also, and more important, tied to academic performance. It is not unreasonable to assume that lower proficiency in English, whether indicated by relatively low TOEFL and GREV scores at application or by a negative recommendation at the end of the first semester, would have a potentially detrimental effect on a student’s ability to cope with academic courses. That assumption is only partially supported by previous studies. In her study of international undergraduate students, Johnson (1988) reported a moderate, yet significant, correlation between entry TOEFL scores and subsequent grade point average (GPA). She also suggested that “the lower the proficiency, the more the role it plays in academic success” (p. 165). However, as Graham (1987) has pointed out, university-level academic programs do not generally admit international students with TOEFL scores in the lowest ranges. As a result, those who might perform poorly (attaining a low GPA) because of very low English-language proficiency (as
evidenced by low TOEFL scores) are never in the equation; therefore, finding a strong correlation becomes inherently less likely. In their study of international graduate students, Light, Xu, and Mossop (1987) reported that TOEFL scores were not effective predictors of academic success, as represented by first semester GPA. Within our population, we were able to access the GPAs of those students \( (n = 59) \) who had completed two years of academic coursework; this permitted us to investigate whether there was any relationship between TOEFL, GREV, or recommendation-type, and the students’ GPA at three intervals, after 12, 18, and 24 months from their entry into the university.

Pearson product-moment correlations between TOEFL and GPA at 12 months \( (r = .24) \), 18 months \( (r = .03) \), and 24 months \( (r = .15) \) were nonsignificant, as were those between GREV and GPA at 12 months \( (r = .20) \), 18 months \( (r = .22) \), and 24 months \( (r = .09) \). The pattern found by Light, Xu, and Mossop (1987) for international graduate students regarding TOEFL and GPA after a single semester was thus confirmed and extended through four academic semesters by our study.

When we looked at the relationship between our positive or negative recommendations and subsequent GPA, we did find one interesting result, illustrated in Figure 8. After 12 months (essentially

---

**FIGURE 8**
Mean GPA for Each Recommendation Group After 12, 18, and 24 Months of Graduate Study

---

PREDICTING SUCCESS FOR INTERNATIONAL TEACHING ASSISTANTS 239
two academic semesters), the mean GPA of those who had received a positive recommendation ($M = 3.61$, $SD = .49$) was significantly higher than that of the negative recommendation group ($M = 3.22$, $SD = .48$), according to an independent samples $t$-test ($t \ [57] = 2.62, p < .01$). This finding does suggest that academic performance during the first year of graduate studies may be affected by the kind of relatively poor ability in spoken English that also leads to the negative recommendation. While the positive group continued to have higher average GPAs than the negative group, by the 18-month point, the difference had narrowed and was no longer significant.

CONCLUSION

We hope that, in reporting the results of this study, we have provided ESL professionals with an empirical basis for some of the arguments that they must present, often in conflict with department heads in other disciplines, concerning the lack of readiness of some international graduate assistants for classroom duties. We also hope that we have shown that a student’s relatively high TOEFL score such as 570 is, by itself, not a guarantee that this individual will convert that general proficiency into a capability to effectively present instructional material in spoken English; for one population, however, it did indicate a substantially better (i.e., twice as likely) chance than did a score of 550. We would advocate that the results of this study be made available to those in other academic disciplines who regularly make long-distance assistantship awards to international students. To avoid the administrative difficulties created by negative recommendations (not to mention the loss of face and other detrimental effects experienced by those students receiving such recommendations), academic administrators might be encouraged, at the very least, to set higher proficiency requirements for those international graduate students to whom they intend to award teaching assignments, while also realizing there is a risk-to-reward ratio, as indicated in our population, associated with most of the proficiency-level scores currently accepted as adequate for graduate school admission.

In interpreting the results of this study, it is important to remember that the recommendation status of the students involved was subsequent to, and in many ways a consequence of, the 15 weeks of training/preparation provided. Without that training, the number of positive recommendations would inevitably have been lower. Even with that period of training, those receiving negative recommendations clearly evidenced ESL needs that had not yet been met. An obvious implication is suggested: If U.S. universities
intend to continue recruiting the brightest and best of the world’s international graduate students to support university teaching and research missions, they should plan to provide more extended periods of adjustment and ESL training for a substantial proportion of those students.

THE AUTHORS

George Yule is Associate Professor in the Linguistics Program and the Communication Disorders Unit of Louisiana State University, where he teaches courses in both theoretical and applied linguistics. He has recently published, with Elaine Tarone, *Focus on the Language Learner* (Oxford University Press, 1989). Earlier works include *The Study of Language* (Cambridge University Press, 1985) and, with Gillian Brown, *Discourse Analysis* (Cambridge University Press, 1983), *Teaching the Spoken Language* (Cambridge University Press, 1983), and *Teaching Talk* (Cambridge University Press, 1984).

Paul Hoffman is Associate Professor in the Communication Disorders Unit of Louisiana State University, specializing in aspects of children’s normal and disordered phonological development. He has recently published, with Gordon Schuckers and Raymond Danilloff, *Children’s Phonetic Disorders: Theory and Treatment* (Little, Brown, 1989).

REFERENCES


Is your program ready for self-study?

When was the last time you evaluated your language program or your professional preparation program? How does your program compare with others? What criteria should you use for evaluation? If you're looking for answers to these questions, you're ready to use TESOL's self-study materials.

TESOL's Committee on Professional Standards has developed an evaluative program that defines useful criteria for self-study and then helps you ask the right questions.

- The TESOL Core Standards for Language and Professional Preparation Programs outlines the attributes of quality programs for ESOL students and professionals undergoing training. These standards will guide your ongoing review. You are asked to endorse these standards and to submit a letter to that affect to TESOL Central Office.

- Having accepted and endorsed the standards of quality programs, you may next purchase TESOL’s Manual for Self-Study and an accompanying booklet appropriate for your program level—elementary and secondary, adult, postsecondary, and professional preparation programs.

TESOL's self-study materials will help you understand your program's strengths and weaknesses and clarify goals.

"The value of conducting a self-study cannot be over-emphasized. We didn't anticipate the extent of information and insights we would gain from applying TESOL's Self-Study program. Self-study is something you will want to integrate into the foundation of your program."*

* John Farneslow, TESOL Program, Teachers College, Columbia University, New York, NY, USA

For more information contact:

Susan Rayley
Director of Field Services
TESOL
1600 Cameron Street, Suite 300
Alexandria, Virginia 22314, USA
(703) 936-0774
The study reported in this article examined the first language and second language reading and writing abilities of adult ESL learners to determine the relationships across languages (L1 and L2) and across modalities (reading and writing) in the acquisition of L2 literacy skills. Specifically, we investigated relationships (a) between literacy skills in a first language and literacy development in a second language (i.e., between reading in L1 and L2, and between writing in L1 and L2), and (b) between reading and writing in L1 and L2 (i.e., between reading and writing in L1, and between reading and writing in L2). The subjects, Japanese and Chinese ESL students in academic settings, were asked to write an essay and to complete a cloze passage in both their first and second languages. The results indicate that literacy skills can transfer across languages, but that the pattern of this transfer varies for the two language groups. It also appears that reading ability transfers more easily from L1 to L2 than does writing ability, and that the relationship between reading and writing skills varies for the two language groups. These data suggest that L2 literacy development is a complex phenomenon for already literate adult second language learners involving variables such as L2 language proficiency, L1 and L2 educational experience, and cultural literacy practices that may be related to different patterns of L2 literacy acquisition.

*This paper was accepted for publication in the TESOL Quarterly under the editorship of Stephen J. Gaies.
Adult language learners have two primary sources from which to construct a second language system: knowledge of their first language and input from the second language. Those adults who are already literate in their first language have these same sources available to them as they develop literacy skills in their second language. They can draw on their literacy skills and knowledge of literacy practices from their first language (interlingual transfer), and they can also utilize the input from literacy activities—reading and writing (intralingual input)—in their developing second language.

There is evidence that second language learners utilize both of these sources in acquiring second language literacy skills. Cummins (1981) makes the strongest case for interlingual transfer of literacy skills. He claims that there is a cognitive/academic proficiency that is common to all languages and that this common language proficiency allows for the transfer of literacy-related skills across languages. According to this interdependence hypothesis,

\[
\text{to the extent that instruction in Lx is effective in promoting \{academic\} proficiency in Lx, transfer of this proficiency to Ly will occur provided there is adequate exposure to Ly (either in school or environment) and adequate motivation to learn Ly. (p. 29)}
\]

Thus, developing literacy in an L2 may be affected by literacy capabilities in the L1.

Some empirical studies have supported Cummins’ claims. Mace-Matluck, Dominguez, Holtzman, and Hoover (1983) studied English literacy among students of Cantonese language background and found a significant correlation between literacy acquired in English and the literacy level achieved in Cantonese prior to English instruction. Goldman, Reyes, and Varnhagen (1984) found that higher order skills involved in comprehending stories in L1 appeared to transfer to comprehension of stories in students’ L2. Edelsky’s (1982) qualitative study of elementary school Spanish-English bilingual children’s writing offered strong evidence of transfer across languages. Canale, Frenette, and Belanger (1988) found a significant relationship between the L1 and L2 writing of French-English bilingual high school students.

Not all studies, however, have presented evidence for such transfer of literacy-related skills. Cummins, for example, suggests that transfer capability emerges only after individuals attain a threshold level of L2 proficiency to permit cognitively demanding language use. Clarke (1978) found evidence that good L1 readers are also good L2 readers. However, his case study of two students, one good and one poor L1 reader of comparable L2 language proficiency, showed that the superiority of the good first language
reader decreased substantially in English. Clarke’s work suggests that some threshold level of language proficiency is necessary for good L1 readers to maintain their good reading skills in the L2. Cziko (1978) also found that only those most proficient in the L2 could take advantage of semantic cues when reading meaningful L2 texts, thus providing additional evidence that a language competence threshold is necessary before transfer of skills can occur. Alderson’s (1984) review of the relationship between L1 and L2 reading ability noted considerable support for the view that a language competence threshold is a prerequisite for the transfer of L1 reading skills.

McLaughlin’s (1987) data also suggest that transfer of literacy skills may not be as automatic as Cummins claims. McLaughlin’s prediction that advanced L2 learners would be better able to utilize effective L1 reading skills and thus be better L2 readers was not supported; rather, his study found that these advanced learners did not perform significantly better in L2 reading than beginning ESL students. His explanation was that advanced L2 learners utilized strategies directed at meaning when they read in their first language, but that they had not yet made this shift in their second language. Because McLaughlin did not test the readers’ first language reading abilities, we cannot be sure that this explanation is valid. Thus, the picture of interlingual transfer of literacy-related skills is complicated by the notion of the language proficiency threshold suggested by Cummins, Clarke, and Cziko, and by the possibility that this threshold may be a necessary but not a sufficient condition for transfer to occur, as McLaughlin suggested.

The second source available to learners acquiring literacy skills in a second language is input from the L2. This intralingual input occurs as the result of literacy events in the second language, which provide the learner with information about the forms, functions, and processes used in literacy activities in the developing language system. Whatever form this second language literacy input may take, it is almost certainly not the case that second language learners acquire reading skills only from reading, or writing skills only from writing. Surveys of first language research (Belanger, 1987; Stotsky, 1983) have shown that there are strong correlations between reading and writing abilities (that is, good readers tend to be good writers and vice versa); these studies have also identified common cognitive processes and/or structural components underlying reading and writing abilities such that instruction and improvement in one skill can enhance capabilities in the other.

In contrast to the substantial body of research on the relationship between reading and writing abilities in a first language, little has
been done to explore this connection for second language learners. Krashen’s (1984) claim that second language learners’ writing competence derives from large amounts of self-motivated reading for interest and/or pleasure remains largely untested and unsubstantiated. Still, it is difficult to imagine that second language input would not play a significant role in developing literacy skills in L2, i.e., reading input presumably affects the development of writing and reading abilities and/or writing input affects the development of reading and writing abilities. The situation for second language learners is much more complex than it is for first language readers and writers: One must take into account not only the learner’s L2 language proficiency, but also the possibility of interaction of first language literacy skills with second language input.

An analysis of second language literacy development, then, must consider both interlingual transfer and intralingual input; it must describe what learners utilize from their first language and what they utilize from second language input as they develop L2 literacy skills. Examination of L2 development must also include analysis of the relationships between literacy skills across languages (from L1 to L2) as well as across modalities (reading to writing and writing to reading).

THE STUDY

Although some research studies (e.g., Canale, Frenette, & Belanger, 1988; Clarke, 1978; Cziko, 1978; McLaughlin, 1987) have looked at the transfer of literacy skills across languages, and a few studies (Janopoulous, 1986; Flahive & Bailey, 1988) have examined reading-writing relationships in L2, no study has yet described how these two strands are related for the same individual engaged in developing literacy skills in a second language. This study was designed to examine the relationships between first and second language reading and writing abilities. By looking at relationships between reading and writing abilities in both first and second language, we can begin to describe the contributions of first language literacy skills and the contributions of second language reading and writing experiences to the development of literacy in a second language. Underlying these issues is the question of the role that language proficiency plays.

We are interested, then, in five basic questions:

(a) What is the relationship between first and second language reading abilities?
(b) What is the relationship between first and second language writing abilities?
(c) What is the relationship between reading and writing in the learner’s first language?
(d) What is the relationship between reading and writing in the learner’s second language?
(e) Does second language proficiency affect interlingual or intralingual transfer?

Method

Subjects. A total of 48 native speakers of Chinese and 57 native speakers of Japanese participated in the study. Of the 105 subjects, 72 were enrolled in preacademic intensive English programs, and 33 were enrolled in basic writing or freshman composition courses as fully matriculated students. The subjects were drawn from four institutions: Georgia State University; Southern Illinois University; the University of California, Los Angeles; and the University of Washington.

All 105 subjects were high school graduates, and almost all (102) had graduated from high school in their native countries. Of the 46 subjects with undergraduate degrees, only 1 had obtained that degree from a U.S. university. In addition, 10 subjects had earned graduate degrees, three of which were obtained at a U.S. university. The average length of time that subjects had been in the U.S. was 11½ months, with a range of 1 month to 17 years.

Subjects’ language proficiency varied from low-intermediate (around 400 TOEFL) to advanced (over 525 TOEFL). Subjects were assigned to one of three language proficiency levels on the basis of either information about their placement in intensive courses or the criteria required for admission into regular curricula. The low-intermediate group (Level 1), which consisted of 8 subjects, was defined as the midrange of intensive programs, with TOEFL scores in the 420-480 range or Michigan Test scores between 42 and 60. The high-intermediate group (Level 2), with 64 subjects, was defined as the highest level(s) of intensive programs, with TOEFL scores in the 480-520 range or Michigan Test scores between 60 and 78. The advanced group (Level 3), with 33 subjects, consisted of those subjects enrolled in regular curriculum classes (i.e., not in intensive programs), with TOEFL scores above 525 or Michigan test scores above 80.

The Chinese, as a group, were approximately the same age as the Japanese ($M = 25.48$ years, $SD = 5.03$, vs. $M = 24.33$ years, $SD = 4.76$), and had studied English approximately the same
amount of time \( M = 8.81 \text{ years}, SD = 3.83, \) for the Chinese vs. \( M = 7.74 \text{ years}, SD = 2.20, \) for the Japanese). Other characteristics of the two groups are shown in Table 1. The Chinese as a group had been in the U.S. longer, with 31% having been in the country more than 12 months as compared to 10% of the Japanese subjects.

**TABLE 1**
Length of Residency in U.S., Estimated L2 Proficiency, and Educational Levels of Subjects

<table>
<thead>
<tr>
<th>Variable</th>
<th>Chinese ((n = 48))</th>
<th>Japanese ((n = 57))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( n )</td>
<td>%</td>
</tr>
<tr>
<td><strong>Months in U.S.</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;6</td>
<td>25</td>
<td>52</td>
</tr>
<tr>
<td>6-12</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td>13-24</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>&gt;24</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td><strong>L2 proficiency</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level 1 (420-480 TOEFL)</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Level 2 (480-520 TOEFL)</td>
<td>19</td>
<td>40</td>
</tr>
<tr>
<td>Level 3 (525+ TOEFL)</td>
<td>25</td>
<td>52</td>
</tr>
<tr>
<td><strong>Highest level of L1 education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Attended high school</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Completed high school</td>
<td>15</td>
<td>31</td>
</tr>
<tr>
<td>Attended university</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>Completed university</td>
<td>18</td>
<td>38</td>
</tr>
<tr>
<td>Attended graduate school</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Completed master's degree</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td><strong>Highest level of L2 education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>24</td>
<td>50</td>
</tr>
<tr>
<td>Attended high school</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Completed high school</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Attended university</td>
<td>16</td>
<td>33</td>
</tr>
<tr>
<td>Completed university</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Attended graduate school</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Completed master's degree</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

The level of education achieved in the first language was nearly equivalent for both groups, with more than half the subjects in each group having completed some university coursework in their first language. Comparing the two groups on L2 academic experience, 50% of the Chinese subjects had attended or were currently attending L2 university classes (not ESL courses), while only 11% of the Japanese subjects had attended or were enrolled in L2 university classes. Furthermore, the Chinese subjects averaged a higher level
of completed or ongoing L2 education. Of the Chinese, 17% (8/48) of the Chinese had either completed a university degree or were pursuing a graduate degree in English, as compared with only 4% (2/57) of the Japanese subjects.

These differences in L2 education are reflected in the language proficiency distribution of Chinese and Japanese subjects. A much higher percentage of Chinese (52% vs. 14% of the Japanese) were functioning at a full proficiency level in English and were not enrolled in intensive English courses. Thus, on three important variables the two groups were not the same: time in the U.S. ($\chi^2[2] = 9.437, p < .01$, comparing the categories less than 6 months, 6-12 months, and more than 12 months in the U.S.), L2 proficiency ($\chi^2[2] = 18.69, p < .001$, comparing the three levels of L2 proficiency), and level of L2 education ($\chi^2[1] = 19.89, p < .001$, comparing those with no L2 educational experience to those having some L2 education).

**Materials.** Materials consisted of writing prompts and cloze passages in both the first and second language. The writing prompts were designed to elicit comparison/contrast rhetorical organization, a common pattern of academic discourse and one that presents a clearly discernible set of tasks. The L1 prompt—designed to be administered in Chinese or Japanese—was translated from the English version (see Appendix A) by native speakers of Chinese and Japanese. Subjects were instructed to discuss choices in career selection based on the relative availability of job options. The L2 prompt—designed to be addressed in English by an ESL population—asked subjects to write about the importance of belonging to a group or of being an individual in order to achieve their goals (see Appendix A). Each prompt was typed at the top of a sheet of paper, which the students could use in writing their essays. Chinese subjects were told that they could write in either complex or simplified characters, and were provided with special composition paper from the People’s Republic of China to accommodate the Chinese characters.

Native speakers of each language were asked to select the reading passages to be clozed using the following criteria: (a) The topic of the passage must be of general interest; (b) the passage must be authentic text aimed at readers with high school level reading skills; (c) the passage must exhibit comparison and contrast rhetorical organization; and (d) the length of the passage must be between 325 and 400 words. Using these criteria, the Chinese language consultant selected an article from a Chinese newspaper (People’s Daily, June/July 1988) about differences in family environments.
The Japanese text, from a volume on international understanding (Nishiyama, 1972), dealt with differences in notions of hierarchy and equality in the U.S. and Japan. The English text discussed the differences in the effects of nature versus nurture in personal development, and was taken from an advanced ESL text (Adams & Dwyer, 1982).

After the passages were selected, they were clozed by native speakers of each language. Because of the difficulty of constructing a rational-deletion cloze test with non-English texts and with nonspecialist consultants, we chose to use a 7th word deletion rate, maintaining the first sentence of each passage intact. Clozing the Chinese and the Japanese texts proved somewhat difficult because there is no clear definition of a word as an orthographic unit in each of these writing systems. The Chinese consultant’s decision was that a word had been identified if a character or characters could stand alone with the same meaning. For example, if the word for sun consisted of the characters for bright and ball, then those two characters were said to constitute one word—sun. In Japanese, an agglutinating language, a word was considered to be more nearly a semantically loaded morpheme. The English passage contained 52 blanks; the Chinese, 44 blanks; and the Japanese, 44 blanks.

Instructions included a sample sentence with words written in the blanks. Although the Chinese text was written in complex characters (the more universally readable by Chinese speakers), the Chinese instructions allowed subjects to write in either complex or simplified characters. All passages were then typed, and the space allotted for each cloze item was standardized across languages.

Procedures. The data were collected in the first month of the fall 1988 academic term. All writing tasks preceded all reading tasks so that the reading passages would not provide models for writing, and thereby affect writing performance. L1 and L2 tasks were counterbalanced.

Subjects were given between 30 and 45 minutes to complete each of the four tasks. Tasks were administered over a 2-week period to ensure that language learning between task administrations would not significantly affect results. No dictionaries were allowed, and subjects were given no additional instructions apart from those appearing with the essay prompts and the cloze passages.

Scoring. Both the first language essays (Chinese and Japanese) and the English essays were evaluated by native speakers of those languages using 6-point scales. Each essay was scored by two raters; essays with scores that differed by two or more scale points were read by a third rater and the extreme score was dropped. The score
for each essay was the average of two raters—either the first two raters, or, in cases in which a third rater was needed, the average of the third rater and the closest score.

The English essays were scored using the 6-point scale developed to score the Test of Written English (TWE) portion of the TOEFL (Appendix B). All three raters had been trained by the Educational Testing Service as TWE raters and had scored TWE exams on several occasions.

Since no guidelines existed for Chinese or Japanese essay scoring, raters of these essays developed a scoring guideline by following a two-step process. First they were asked to sort the essays into six piles, with each pile corresponding to a degree of proficiency: Essays ranked 6 were the best, and 1 the worst. Then the raters were asked to write a set of descriptors characterizing the features of each of the six groups of essays, resulting in a written 6-point scale for Chinese (Appendix C) and Japanese (Appendix D) essays. Since criteria defining writing proficiency are likely to be language-specific, we did not expect the 6-point scales developed for the Chinese and the Japanese essays to be equivalent to each other, nor to the 6-point scale used to rate the English essays. Nevertheless, some of the descriptors are surprisingly similar (see Appendices B, C, and D), particularly those focusing on coherence, topic development, and language usage.

The Chinese and Japanese raters were selected according to the following criteria: (a) They had to be familiar with academic writing in their native language, and (b) they had to be fairly sophisticated about language usage. Raters for the Chinese essays were both graduates of Chinese universities. One was a visiting scholar and ESL teacher from the People’s Republic of China; the other was a Chinese language teacher from Taiwan. Raters for the Japanese essays were both graduates of Japanese universities: One was a Japanese language instructor, the other an experienced EFL teacher enrolled in a MATESL program.

Estimates of interrater reliability (coefficient alpha) for the two primary raters in each essay category are reported in Table 2, along with percent of rater agreement, rater means, and standard deviations. (Coefficient alpha is a generalization of the Kuder-Richardson 20 internal consistency reliability estimate for nondichotomously scored items such as essay scores [Shrout & Fleiss, 1979].) Although a third rater was used to provide as accurate an average holistic rating as possible for use in analyses, coefficient alpha and percent agreement are reported to provide information about the functioning of the three 6-point holistic scoring scales. Rater agreement was operationally defined as ratings within one
scale point. Coefficient alphas ranged from .73 (L2 essay) to .91 (Japanese essay), and are similar to the .85-.88 range reported in the *Test of Written English Guide* (Educational Testing Service, 1989, p. 12). The .73 alpha reported for L2 essay raters is low, in part due to the relatively restricted variability of the second rater’s ratings ($SD = .77$). The agreement rate between the two L2 essay raters of 92% and the reported alpha are evidence of rating system reliability. Reliability estimates can be boosted in the continuation of this work by adding raters or adding a wider range of L2 essays. In this sample, either one or both of the raters assigned ratings of 3 to 71% of the L2 essays and 4 to 84%.

### TABLE 2
- **Essay Rater Means, Standard Deviations, Percent Agreement, and Coefficient Alphas**

<table>
<thead>
<tr>
<th>Essay language</th>
<th>Rater</th>
<th>Percent agreement</th>
<th>Coefficient alphas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Chinese ($n = 48$)</td>
<td>3.52</td>
<td>2.50</td>
<td>.67</td>
</tr>
<tr>
<td>$M$</td>
<td>1.69</td>
<td>1.29</td>
<td></td>
</tr>
<tr>
<td>Japanese ($n = 57$)</td>
<td>3.18</td>
<td>3.19</td>
<td>.97</td>
</tr>
<tr>
<td>$M$</td>
<td>1.63</td>
<td>1.46</td>
<td></td>
</tr>
<tr>
<td>English ($n = 105$)</td>
<td>3.48</td>
<td>3.27</td>
<td>.92</td>
</tr>
<tr>
<td>$M$</td>
<td>1.03</td>
<td>.77</td>
<td></td>
</tr>
<tr>
<td>$SD$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Because the sample was initially distributed into six categories by raters, the method used for constructing the L1 essay rating scales yielded greater rating variability compared to the L2 scales. Coefficient alpha for the Chinese essays was .83, although the rater agreement was only 67%. The first Chinese rater consistently rated essays higher than did the second rater. The Japanese essay ratings had both higher interrater reliability (.91) and higher rater agreement (97%). Future research efforts will address the issues of L1 and L2 scale equivalency and rating variability raised in this exploratory study.

Cloze passages were scored using exact-word scoring, since Oller’s (1979) review of cloze research indicated that although percentage scores may be lower with exact-word scoring, rank order should remain the same with exact-word or acceptable-substitute scoring.

254 TESOL QUARTERLY
RESULTS

As we have seen, the two subject groups differed on three important variables: time spent in the U. S., level of L2 education, and level of L2 proficiency. These differences in background should be kept in mind as possible sources of influence on the analyses relating L1 and L2 language skills.

Mean scores by task are reported in Table 3. The mean for the Chinese cloze test was 10.9, and for Japanese the mean was 32.8, out of a total of 44 blanks on each test. This difference may well reflect the different procedures used to cloze the texts, with the Japanese text possibly being more predictable than the Chinese. In any case, since we were interested in rank order for each language group, and since we did not require text comparability across languages, these differences in means were not of concern. The mean score of the Chinese subjects on the English cloze was 14.94; the Japanese mean was 11.04 (52 blanks total), reflecting the different L2 language proficiencies of the two groups.

<table>
<thead>
<tr>
<th>Task</th>
<th>Maximum possible score</th>
<th>Chinese (n = 48)</th>
<th>Japanese (n = 57)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>L1 cloze</td>
<td></td>
<td>10.90</td>
<td>4.14</td>
</tr>
<tr>
<td>L2 cloze</td>
<td></td>
<td>14.94</td>
<td>6.08</td>
</tr>
<tr>
<td>L1 essay</td>
<td></td>
<td>3.21</td>
<td>1.56</td>
</tr>
<tr>
<td>L2 essay</td>
<td></td>
<td>3.55</td>
<td>.75</td>
</tr>
</tbody>
</table>

The L1 essay scores (Table 3) were comparable for the two groups: The Chinese mean was 3.21, the Japanese, 3.25. The English essay scores again reflected the difference in L2 language proficiency: The Chinese mean rating was 3.55, the Japanese, 3.07.

Relationships between L1 and L2 reading and writing were investigated initially by examining correlation coefficients. Weak to moderate correlations are reported in Table 4. Other research (Canale, Frenette, & Belanger, 1988) has also reported weak to moderate correlations between L1 and L2 writing abilities. Correlation magnitudes for the reading-writing relationship may also be considered in terms of Shanahan and Lomax’s (1986) proposed model of the reading-writing relationship, which argues...
for the existence of multiple relations (i.e., the interactions among language skills such as word analysis, spelling knowledge, and word recognition may differ within and across discourse levels), as well as the possibility that the nature of the reading-writing relationship might change with development and thus not be linearly related. In this case, the Pearson correlation thus may underestimate the actual relationship between these two language skills.

TABLE 4
Correlations by Language Groups for L1 and L2 Reading and Writing Tasks

<table>
<thead>
<tr>
<th></th>
<th>Chinese (n = 48)</th>
<th>Japanese (n = 57)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1 reading × L2 reading</td>
<td>( r = .366^{**} )</td>
<td>( r = .509^{**} )</td>
</tr>
<tr>
<td>L1 writing × L2 writing</td>
<td>( r = -.019 )</td>
<td>( r = .290^{*} )</td>
</tr>
<tr>
<td>L1 reading × L1 writing</td>
<td>( r = .271^{*} )</td>
<td>( r = .403^{**} )</td>
</tr>
<tr>
<td>L2 reading × L2 writing</td>
<td>( r = .494^{**} )</td>
<td>( r = .271^{*} )</td>
</tr>
</tbody>
</table>

* \( p < .05 \)
** \( p < .01 \)

The correlations in Table 4 show the following relationships: (1) L1 and L2 reading scores had weak to moderate correlations for both the Chinese and the Japanese subjects; (2) L1 and L2 writing scores showed a weak positive correlation for the Japanese but not for the Chinese subjects; (3) L1 reading and writing showed weak to moderate correlations for both groups, as did L2 reading and writing. For both groups, there are stronger relationships between reading abilities across languages than between writing abilities across languages. The L1-L2 writing relationship for Japanese is weak, and for Chinese is not significant. The correlations in Table 4 also show that for the Chinese the relationship between reading and writing is strongest in L2, but for the Japanese the reading-writing relationship is strongest in L1. Reading and writing are related, but the strength and nature of the relationship differs for each of these language groups, either due to language or other background variable differences.

The means and correlations by L2 proficiency levels also showed pattern differences by language groups and by levels. However, because of the \( n \)-sizes for Level 1 (4 each for Chinese and Japanese) and for the Level 3 Japanese group (8), it is difficult to draw any conclusions about the role that language proficiency plays in these reading and writing relationships. Two trends, although not significant, did appear that are worth noting. First, as we might
expect, L2 reading and writing scores tended to increase as L2 proficiency increased. Second, however, L1 reading and writing scores tended to decrease as L2 proficiency increased. This was particularly noticeable for L1 writing, where means for both groups decreased from an average of 3.38 at Level 2 proficiency, to 2.86 at Level 3. It appears that L1 writing skills are rated weaker as L2 proficiency increases. This apparent decrease in writing ability may reflect the fact that students in an L2 academic environment are not engaging in L1 academic writing activities of the type we were measuring. The resulting attrition may be similar, then, to the phenomenon of language loss that occurs when a language is no longer used sufficiently to maintain proficiency. Still, these results by language proficiency must be interpreted cautiously, given the disparate n-sizes of the levels.

Partial correlations were calculated to further examine the relationships between L1 and L2 skills and L1 and L2 educational experiences. For Chinese, when L1 educational experience was statistically controlled for (using partial correlations, with L1 education categories coded 0-6), the small but significant \( r = .271, p = .031 \) correlation between L1 reading and writing disappeared \( (r = .073, p = .314) \), suggesting that L1 educational experience may be an intervening variable causally linking L1 reading and writing. In other words, the relationship between L1 reading and writing for Chinese depends upon educational experience level in L1. However, the relationship for Chinese subjects between L2 reading and writing \( (r = .494, p = .000) \) remained, even after all measures of L2 input (L2 education level, total English studied, and months in the U.S.) were controlled for through partial correlation. This suggests that, for Chinese subjects, reading and writing are more directly related in L2 than in L1.

For the Japanese, unlike the Chinese, controlling for L1 educational experience did not affect the correlation between L1 reading and writing \( (r = .493, p = .000) \). Like the Chinese, controlling for L2 educational experience, total English, and months in the U.S. did not affect the correlation between L2 reading and writing for the Japanese \( (r = .271, p = .021) \).

Stepwise multiple regressions for both L1 and L2 reading and writing scores were used to determine the best subset of variables that predict those scores. Variables used in the procedure included the following: L1 and L2 reading and writing scores, level of L1 and L2 education, length of time (months) in the U.S., and amount (total months) of English language study reported. Because the study was exploratory in nature, the significance level for variable elimination was set at .10; this permitted us to capture the significant
relationships as well as to identify those variables that might be explored more fully in future studies. The adjusted R's reported in Table 5 represent the proportion of variance in the dependent variable (L1/L2 cloze, L1/L2 essay) accounted for by the other variables in the equation. Cohen (1977) had suggested that adjusted R's used as estimates of effect size in behavioral sciences research are considered small in the .02-.12 range, medium in the .13-.25 range, and large if .26 or greater. Applying Cohen's criteria, the adjusted R's reported in Table 5 are all in the medium and large range.

### TABLE 5
Results of Stepwise Multiple Regression by Language Group for L1 and L2 Reading and Writing Tasks

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>R² (adjusted)</th>
<th>Independent variable</th>
<th>Standardized regression coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1 cloze</td>
<td>.369</td>
<td>L2 cloze</td>
<td>.230*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L1 education</td>
<td>.529**</td>
</tr>
<tr>
<td>L2 cloze</td>
<td>.431</td>
<td>L1 cloze</td>
<td>.290*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L2 essay</td>
<td>.461**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L2 education</td>
<td>.262**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L1 education</td>
<td>.242*</td>
</tr>
<tr>
<td>L1 essay</td>
<td>.160</td>
<td>L2 education</td>
<td>-.247*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L1 education</td>
<td>.335**</td>
</tr>
<tr>
<td>L2 essay</td>
<td>.355</td>
<td>L2 cloze</td>
<td>.591**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L1 education</td>
<td>-.384**</td>
</tr>
</tbody>
</table>

#### Chinese (n = 48)

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>R² (adjusted)</th>
<th>Independent variable</th>
<th>Standardized regression coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1 cloze</td>
<td>.386</td>
<td>L2 cloze</td>
<td>.417**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L1 essay</td>
<td>.396**</td>
</tr>
<tr>
<td>L2 cloze</td>
<td>.284</td>
<td>L1 cloze</td>
<td>.508**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total English</td>
<td>.229*</td>
</tr>
<tr>
<td>L1 essay</td>
<td>.229</td>
<td>L1 cloze</td>
<td>.492**</td>
</tr>
<tr>
<td>L2 essay</td>
<td>.186</td>
<td>In U.S.</td>
<td>.273**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L1 education</td>
<td>.331**</td>
</tr>
</tbody>
</table>

#### Japanese (n = 57)

* *p < .10
** *p < .05

258 TESOL QUARTERLY
As the results show, for both Chinese and Japanese, L2 cloze is a significant variable for L1 cloze, and L1 cloze is a significant variable for L2 cloze, as we would expect from the L1-L2 reading correlations. This L1-L2 reading relationship, however, is the only relationship that the two language groups share; except for this, for Chinese and Japanese, different variables predict task performance. On the L1 cloze task for the Japanese (but not the Chinese), L1 essay scores contribute significantly to prediction of L1 cloze scores, reflecting the L1 reading-writing correlation we have already noted for the Japanese. On the L2 cloze task for the Chinese (but not the Japanese), L2 essay scores contribute significantly to prediction of L2 cloze scores, reflecting the L2 reading-writing correlation we have already noted for the Chinese.

The interesting finding of this analysis is the prevalence of L1 and/or L2 education as a predictor of task performance for Chinese subjects. For L1 cloze and L2 cloze, the relationship between level of education and reading performance is positive. That is, the higher the level of education either in L1 or L2, the better the reading scores.

However, for the writing tasks, the relationships are different. For the L1 essay, the relationship between L1 education and L1 essay scores is positive—the higher the level of L1 education, the better the L1 writing score. However, the relationship between L2 education and L1 essays is negative—the higher the level of L2 education, the worse the L1 writing score. (Or, conversely, the lower the level of L2 education, the better the L1 writing score.) The situation is similar for L2 essay scores. The lower the level of L1 education, the higher the L2 writing score. Or, the higher the level of L1 education, the worse the L2 writing score. For the Japanese, however, L1 education is a variable that is positively related to L2 essay scores. That is, the higher the level of L1 education, the better the L2 essay.

DISCUSSION

The data suggest that interlingual transfer can occur, but that the pattern and the strength of this transfer varies according to first language background and other aspects of educational background and experience. For reading, the transfer from L1 to L2 was similar for Chinese and Japanese, but for writing, the transfer from L1 to L2 was different. These differences may be a function of L2 language proficiency. Another possibility is that real cultural differences are reflected in the literacy practices and abilities of the two groups.

READING-WRITING RELATIONSHIPS IN L1 AND L2
The results also suggest that reading ability transfers more easily from L1 to L2 than does writing ability. In addition to the weak relationship noted in the L1-L2 writing correlations for both groups, the multiple regression analyses indicate that although reading scores predict reading scores in either language for both groups, writing never appears as a variable that predicts writing. Again, these differences may be the result of L2 language proficiency or of cultural variables. Further research is needed to determine whether the different variables that predict Chinese and Japanese writing scores are the result of L2 proficiency differences or whether these variables reflect real cultural differences in writing skills. It appears that L1 and L2 educational levels are implicated in complex ways.

Like interlingual transfer, intralingual input (the relationship between reading and writing skills within a language) did not show the same patterns for Chinese and Japanese learners. L1 reading-writing relationships were apparently correlated for both Japanese and Chinese, but as we saw in the partial correlations, for the Chinese, L1 reading-writing relationships may be accounted for by differences in L1 educational experience. For the Japanese, the L1 reading-writing relationship remained significant even when L1 education was taken into account. At the same time, the Chinese showed a stronger correlation between L2 reading and writing abilities than did the Japanese.

PEDAGOGICAL IMPLICATIONS

Given the exploratory nature of this study, any teaching implications based on these preliminary findings should be treated with caution. Still, the results suggest some general implications for the classroom.

The significant correlations between L1 and L2 reading for both groups shows that there is a relationship between reading in the first and reading in the second language. While the correlation was significant, the two are not perfectly correlated. Although other factors are important, the relationship could and should be exploited in L2 reading pedagogy. L1 reading skills can and should be used in L2 reading pedagogy, but the wise instructor will not and should not depend on an automatic transfer of L1 reading abilities/skills to L2 reading.

The weak correlation between L1 and L2 writing for the Japanese, and the lack of correlation of L1 and L2 writing for the Chinese (whose proficiency was higher), suggest that the extent to which L1 writing may be exploited or used in L2 writing pedagogy
may be limited to lower L2 proficiency levels and/or certain L1 language groups. Hence, the writing teacher may rely even less than the reading teacher on the transfer of L1 writing skills to L2 writing.

The differences in the reading-writing relationships between the Japanese and Chinese groups suggest that if the nature of the L1 and L2 reading-writing relationship changes as L2 proficiency develops, then the extent to which L1 may be relied on in L2 pedagogy also changes with L2 literacy development. That is, whereas teachers may be able to exploit L1 literacy relationships in the transfer of L2 literacy practices at lower proficiency levels, they cannot do so reliably at more advanced L2 levels. At more advanced L2 levels, teachers need to rely more on the developing L2 literacy. In other words, at lower proficiency levels, interlingual transfer may be more important, whereas at higher proficiency levels, intralingual input may be the more significant source for developing L2 literacy skills.

**CONCLUSION**

The data from this study suggest that L2 literacy development is a complex phenomenon for already literate adult L2 learners. However, this study is exploratory in nature and raises several new questions. In particular, we need to (a) examine carefully L2 language proficiency as a variable in literacy development; (b) learn more about the relationship between L1 and L2 writing, and about the ways in which L2 writing skills are affected by interlingual transfer and intralingual input; and (c) investigate further the literacy practices of these two groups that may relate to different patterns of L2 literacy acquisition.

**ACKNOWLEDGMENTS**

This is a revised version of a paper presented at the 23rd Annual TESOL Convention in San Antonio, TX, March 1989. We thank former TESOL Quarterly Editor Stephen J. Gaiés for his editorial assistance.

A research project such as this is an enormous undertaking and would have been impossible without the help of the following people, who provided invaluable assistance with materials preparation, data collection, and scoring: at Northwestern University, Zaida Cintron and Pilar Esteves; at Georgia State University, Becky
Bodnar, Alice Gertzman, Sissy Kegley, Wendy Newstetter, Karen Peterson, and Mei-En Wang; at Southern Illinois University, Ming Wang and Maria Linz, as well as the staff at the Center for English as a Second Language, and instructors of Freshman Composition for international students; at the University of Washington, the students, staff, and administration of the English as a Second Language Programs, especially Patricia Grogan, and Stuart Aque, Kaeko Inoue, Motoo Kobayashi, Hiroshi Matsumoto, Tamie Ninomiya, and Samuel M. Shepherd; at the University of California, Los Angeles, Jan Frodesen, Christine Holten, and Juan Carlos Gallegas; at San Diego State University, Bill Gaskill. In addition, we would like to acknowledge research assistance and funding from the Research Office at Georgia State University, from the Office of Research Development and Administration at Southern Illinois University, and from the English Department and the Graduate School Research Fund at the University of Washington.

THE AUTHORS

Joan Eisterhold Carson is Assistant Professor in the Department of Applied Linguistics and ESL at Georgia State University, where she teaches graduate courses. Her research interests center on reading-writing relationships, language attrition, and biliteracy.

Patricia L. Carrell is Dean of the Graduate School and Professor of English at the University of Akron in Ohio. Her research on schema theory and second language reading has been widely published in refereed journals and in two recently coedited books, Research in Reading in English as a Second Language (TESOL, 1987) and Interactive Approaches in Second Language Reading (Cambridge, 1988).

Sandra Silberstein is Associate Professor of English at the University of Washington, where she teaches graduate courses in the MATESL program, as well as courses on language variation, and narrative and critical theory. She is coauthor of Reader's Choice (University of Michigan Press, 1977/1988) and the forthcoming Techniques in Teaching Reading (Oxford University Press). Her research interests include second language reading, narrative and critical theory, ethnicity, and gender.

Barbara Kroll is Assistant Professor of English at California State University, Northridge, where she teaches courses in linguistics, and trains teaching assistants for the first-year composition program. She recently edited Second Language Writing: Research Insights for the Classroom (Cambridge University Press, 1990). She recently completed five years as a consultant for the Educational Testing Service, serving on the test development committee for the Test of Written English.

Phyllis A. Kuehn is Assistant Professor in the Department of Applied Linguistics and ESL at Georgia State University, where she directs the testing program and teaches graduate courses. Her research interests are language assessment, and employment and certification testing. She will be joining the faculty of the School of Education and Human Development at California State University, Fresno in the fall of 1990.
REFERENCES


READING-WRITING RELATIONSHIPS IN L1 AND L2 263


**APPENDIX A**

**Writing Prompts Topics for Writing Samples**

*L1 (Translated into Chinese and Japanese)*

As certain professions become more crowded, students of today are faced with a difficult choice. Should they choose a field they wish to enter even though job prospects are limited? Or should they choose a field in which they are less interested but which offers more job opportunities? Which decision would you make? Give reasons for your answer.

*ESL*

Some people believe that it is important for them to be part of a group in order to achieve their goals. Others believe that they should try to be unique individuals in order to reach their objectives. Which position do you prefer? Give examples to support your choice.

**APPENDIX B**

**Test of Written English (TWE) Scoring Guidelines**

6 Clearly demonstrates competence in writing on both the rhetorical and syntactic levels, though it may have occasional errors.

A paper in this category
- is well organized and well developed
- effectively addresses the writing task
- uses appropriate details to support a thesis or illustrate ideas
- shows unity, coherence, and progression
- displays consistent facility in the use of language
- demonstrates syntactic variety and appropriate word choice

5 Demonstrates competence in writing on both the rhetorical and syntactic levels, though it will have occasional errors.

A paper in this category
- is generally well organized and well developed, though it may have fewer details than does a 6 paper
- may address some parts of the task more effectively than others
- shows unity, coherence, and progression
- demonstrates some syntactic variety and range of vocabulary
- displays facility in language, though it may have more errors than does a 6 paper

4 Demonstrates minimal competence in writing on both the rhetorical and syntactic levels

A paper in this category
- is adequately organized
- addresses the writing topic adequately but may slight parts of the task

--uses some details to support a thesis or illustrate ideas
-demonstrates adequate but undistinguished or inconsistent facility with syntax and usage
-may contain some serious errors that occasionally obscure meaning

3 Demonstrates some developing competence in writing, but it remains flawed on either the rhetorical or syntactic level, or both.
A paper in this category may reveal one or more of the following weaknesses:
- inadequate organization or development
- failure to support or illustrate generalizations with appropriate or sufficient detail
- an accumulation of errors in sentence structure and/or usage
- a noticeably inappropriate choice of words or word forms

2 Suggests incompetence in writing.
A paper in this category is seriously flawed by one or more of the following weaknesses:
- failure to organize or develop
- little or no detail, or irrelevant specifics
- serious and frequent errors in usage or sentence structure
- serious problems with focus

1 Demonstrates incompetence in writing.
A paper in this category will contain serious and persistent writing errors, may be illogical or incoherent, or may reveal the writer’s inability to comprehend the question. A paper that is severely underdeveloped also falls into this category.

APPENDIX C

Evaluation Scale Descriptor Chinese Essay

6 The essay is well written, characterized by thoughtful and coherent reasoning.
The essay plan is clearly signaled by transitions.
The overall presentation of argument is convincing, with varied sentence constructions and persuading evidence.
The main idea is identified or implied.
Superior control of language.

5 A clear understanding of the topic is demonstrated.
The argument is unified and coherent with the subject.
The essay plan is logical.
Opening and closing statements are related to each other.
Ideas are sufficiently developed.
There may be some minor errors in usage and sentence structure.

4 The subject is clear.
Some sequence of ideas.
The essay gives directions to subsequent reasoning.
The essay completes the basic task of the assignment.
Not enough convincing evidence to support the main point.
Some off-topic sentences.

3 The subject is identified.
The main idea is stated.
The essay plan is logical.
Reasoning is not adequate or convincing.
No exhaustive argument.

2 The essay plan is presented.
Little development of ideas.
The main point is not clear.

READING-WRITING RELATIONSHIPS IN L1 AND L2 265
APPENDIX D

Evaluation Scale Descriptor Japanese Essay

6 The argument presented is very clear.
   The sequencing of words and sentences is consistent and smooth.
   The topic is addressed well.
   The overall presentation is well organized.
   The vocabulary is abundant.

5 The argument is clear.
   The persuasion is a little weaker than the level 6.
   The fluency of the language is good.
   The vocabulary used is not as elaborate as that in the papers of the level 6.

4 The overall control of the language is more than the average, but not completely satisfactory.
   The argument mostly follows the topic.
   The variety and the type of sentence construction used need more consideration.

3 The argumentation, sequencing of the sentences, expression and vocabulary are acceptable.
   The level is average.

2 The logical development is missing.
   The argument is not clear.
   Some papers are too casual for an essay.
   The vocabulary used in papers is limited.
   The overall length of the papers is too short to develop the argument.

1 The topic is not addressed well.
   The statements are off the point.
   Some of the students misunderstand the question.
   The papers lack the clear arguments about the topic.
The study reported here investigates the relationship between overall comprehension and the comprehension of coreferential pronouns for second language readers of English. In the first phase of the study, L2 students at The Ohio State University read a passage of contemporary U.S. literature. Overall comprehension was measured by an immediate recall protocol, and coreferent comprehension was measured by a coreferent-identification task. In the second phase of the study, both L1 and L2 subjects were observed. Discriminant function analysis, along with information from interviews, provided insight into the types of errors made by L1 and L2 readers. The results suggest that misunderstanding of coreferential ties reflects a misunderstanding of the descriptive phrases to which the pronouns refer.

Research has indicated that understanding how a passage is structured is an important factor in reading comprehension (Bartlett, 1978; Mandler & Johnson, 1977; Meyer, 1979, 1982; Meyer, Brandt, & Bluth, 1980; Rumelhart, 1975; Stein, 1979). More specifically, it has been clearly established that readers who are able to identify top-level structure (“the rhetorical relationship that can interrelate the greatest amount of text” [Meyer, 1985, p. 269]) as well as paragraph-level relationships of a passage are better able to understand the passage than are those readers who remember only a collection of details. An area of structure comprehension that has remained relatively unexplored, however, is that of microstructure—that aspect of text structure that refers to the author’s progression from given to new information. Included in the structure at this level is the use of certain linguistic cues that provide
the reader with information about what the author expects the reader to be holding in memory. Of particular interest in the present study is the comprehension of coreferential pronouns (e.g., he, she, they), which are anaphoric in nature. That is, they are interpretable only through relationships that are presupposed in the text. (Anaphoric relationships as defined by Bosch [1983] and Kameyama [1985] refer to either preceding or subsequent entities in the text.) Because an author designates concepts in a manner deemed appropriate for an L1 reader, it is likely that comprehension of coreferential ties would place a greater cognitive burden on a second language reader than on a first language reader. The primary focus of this study, then, was to explore the role of coreferential pronouns in discourse and the relationship of coreferential-tie comprehension to overall reading comprehension.

Recent research in linguistics and reading comprehension recognizes that language has a dynamic quality: Written text is constructed so as to proceed from given to new information. Reading comprehension depends, in large measure, on the extent to which concept designation, in the text, has been felicitous for a particular reader (Bloom & Hays, 1978). The Lakoff-Cole scale (Cole, 1974, citing and extending by one step Lakoff, 1968) ranks designation of concepts from most to least informative:

- **Proper name:**
  *George Adams* is on the phone again.

- **Definite description:**
  *The furniture salesman* is on the phone again.

- **Epithet:**
  *That jerk* is on the phone again.

- **Pronoun:**
  *He* is on the phone again.

- **Zero:**
  George Adams wants [Ø] to see you.

According to Bloom and Hays (1978), there are two general rules for designation:

1. A designating expression must be chosen that arouses the most understanding in a given audience.
2. The expression that designates a concept on a given occasion must be the least informative that could designate it successfully.

That pronouns are ranked low on the Lakoff-Cole scale (i.e., that they contain relatively little information) suggests that they are used to refer to concepts that the author assumes are readily accessible to the reader. Therefore, the reader’s ability to link a pronoun with the
concept referred to by the author is a critical component of the reading process.

There is potential for communication difficulties for second language readers with respect to pronoun comprehension. A text will likely designate concepts with “the least informative adequate expression” (Bloom & Hays, 1978, p. 32) deemed appropriate for an L1 reader. Because L2 readers are likely to interpret those concepts on the basis of cultural and literacy experiences in another language, comprehension may break down. An L2 reader with the necessary linguistic competence to understand “the ticket salesman is on the phone again,” for example, may or may not have the necessary knowledge of the target culture (that telephone sales are a common practice in U.S. culture, that they occur in private homes as well as offices, that they frequently occur at inconvenient times) to understand the phrase as well as would an L1 reader. Because languages differ with respect to level of demand made upon the reader/hearer in linking pronouns with discourse entities (Huang, 1984; Kameyama, 1985), L2 readers might approach the task of interpreting coreferential pronouns with expectations different from those of an L1 reader. In other words, the likelihood is increased that an L2 reader would encounter difficulties not encountered by an L1 reader.

It is possible that L2 readers’ expectations of pronouns may be guided by characteristic usage in their first language. Huang (1984) and Kameyama (1985) have suggested that languages can be classified according to the extent to which they allow the use of a zero pronoun. Kameyama posits that pronoun usage in all languages is controlled by both syntactic and pragmatic constraints. She refers to the syntactic constraint as a “syntactic overtness requirement” (SOR, p. 7), the extent to which pronouns must be overtly stated, and the pragmatic constraint as “zero anaphora permissibility” (ZAP, p. 3), the extent to which pronouns may be omitted if the information they convey is easily recoverable. Kameyama claims that “SOR and ZAP are ‘two sides of a coin’ in that the absence of SOR simply means the maximal ZAP and the minimal ZAP corresponds to the strongest SOR” (pp. 7-8). According to her analysis, the grammatical systems of those languages with strong SOR “must have overt subject and object phrases even when their reference is conceptually redundant” (p. 6). The grammatical systems of languages without strong SOR allow zero anaphora “whenever the reference is immediately recoverable in discourse” (p. 8). Kameyama’s ZAP typology can be summarized as follows: Languages such as English and French have a strong SOR and therefore permit zero pronouns only under very narrow circumstances, such as in...
I would like [Ø] to see you tomorrow.

Moderate ZAP languages (e.g., Spanish and Arabic) are those languages in which pronouns may be eliminated when the information that they convey is contained elsewhere in the text (such as in verb inflection).

Did you see Craig?
Si, lo vi.
Yes, him [Ø] saw.
Yes, I saw him.

Examples of maximal ZAP languages are languages such as Chinese, Japanese, and Korean. These languages would allow zero pronoun usage in exchanges such as the following:

Did you see Craig?
Yes, [Ø] saw [Ø].
Yes, I saw him.

What these differences imply is that L2 readers might differ in the degree to which they expect the topic of a discourse to guide pronoun usage. Speakers of maximal ZAP languages seem to find pronouns unnecessary as long as the topic does not change. When the main topic or even a detail topic changes, pronouns must be used to indicate that change (Hinds, 1978, pp. 165-166). English, a SOR-dominated language, always demands grammatical subjects and frequently uses pronouns to fill these “slots.” Therefore, there is a less reliable relationship between pronoun usage and change of topic. It was hypothesized that L2 readers accustomed to one type of system might have difficulty adjusting to the target language system.

In sum, recent views of pronoun usage suggest that pronouns provide important links in the “conceptual chain” comprising a message. Because coreferential pronouns, by their very nature, convey a limited amount of meaning and must presuppose another entity, their use makes greater cognitive demands on the reader than would a designator higher on the Lakoff-Cole scale. Of interest in the present study was how readers from different L1 anaphoric systems and proficiency levels interpret pronouns in passages designed for a native audience.

THE STUDY

The main objective of the study reported in this article was to
investigate the relationship between coreferential-tie comprehen-
sion and overall reading comprehension for readers of English as a
second language (L2 readers). In the first phase of this study,
specific information was sought through analyses of the relationships between (a) proficiency level and comprehension of corefer-
tential ties, (b) proficiency level and overall reading comprehension,
(c) L1 anaphoric system and comprehension of coreferential ties,
and (d) L1 anaphoric system and overall reading comprehension.
The purpose of the second phase of this study, which consisted of
the application of the same procedures to L1 readers, was (a) to
observe and compare reading patterns of L1 readers with those of
L2 readers and (b) to investigate the relationship between overall
comprehension and comprehension of coreferential ties for both L1
and L2 readers.

Subjects

Subjects for the first study were 31 international students enrolled
in the English language programs at The Ohio State University.
Subject classification was determined by matriculation status. The
“prematriculation” group consisted of subjects whose TOEFL
scores were less than the required 500 for entrance to the University.
These were intermediate-level students in a program designed to
aid students in (a) adjusting to U.S. culture and (b) developing
sufficient English skills to be able to perform adequately in an
academic setting. Their TOEFL scores ranged from 400 to 450. This
prematriculation level was selected because, by this level, the
students have attained reading speeds of at least 150 words per
minute and received instruction in reading coreferential ties.

Matriculated subjects were volunteers from the intermediate
level of the ESL writing program at the University. This group of
subjects represented those international students who had met the
language requirement for the University yet needed to further
develop their writing skills. Their TOEFL scores ranged from 600
to 650.

Subjects at each proficiency level represented two L1 anaphoric
systems: “moderate ZAP” and “maximal ZAP.” Language back-
grounds represented in the “moderate ZAP” group were Spanish,
Arabic, Portuguese, and Italian. Language backgrounds repre-
sented in the “maximal ZAP” group were Korean, Malaysian,
Chinese, and Japanese. Numbers of subjects per cell were as
follows: prematriculation/moderate ZAP—6; prematriculation/
maximal ZAP—7; matriculated/moderate ZAP—7; matriculated/
maximal ZAP—11.
Subjects were paid volunteers who were recruited by the experimenter through class visitation. Because there were so few students in the prematriculation and matriculation groups, all volunteers \((N = 31)\) were included in the investigation.

For the second phase of this study, L1 subjects were recruited from The Ohio State University first-year composition classes. A randomized sampling of these students produced 28 subjects (a number approximately equal to that of the L2 subjects).

Materials

The text—the introduction to Joyce Carol Oates’ "Demons" (1970)—was a 327-word literary passage in which pronouns are used in a deliberate fashion in order to convey a message (see Appendix A). The text takes the form of a narrative monologue in which the reader views the protagonist’s world through the eyes of the “experience,” the protagonist herself. A literary passage was used because the experimenter found that literary texts, in contrast to expository prose, are more likely to contain large numbers of diverse coreferential ties, including some that must be inferred. This passage was selected because it contains a minimal amount of cultural bias and a high density of coreferential pronouns. The passage was judged to contain a minimum of cultural bias because the event described—a woman walking a dog in the contemporary United States—was considered to be one that had probably been observed by most international students living in the U.S. (The extent to which all texts are, to some degree, culturally biased, must remain the subject of future studies.) Because it was anticipated that subjects from some cultures might have difficulty understanding the character’s thoughts (the young woman is feeling hostile toward her father), all subjects were told that the passage contained an expression of a person’s innermost thoughts.

Twenty-nine coreferential pronouns in the passage were selected for testing. Their selection was based on (a) the purpose of the investigation and (b) the results of a pilot study. *Both* is not, technically speaking, a coreferential pronoun, but, because of its usage with ellipsis in the passage, it functions anaphorically and was therefore included in the coreferent-identification task. *Itself* as a reflexive pronoun bears a syntactic relation rather than an anaphoric relation with its antecedent (Bosch, 1983), but its inclusion was based on an interest in determining whether L2 readers understood even “obvious” ties. Finally, results of pilot study interviews indicated that many subjects who understood the meaning and usage of *this* in the phrase, *this dog*, were unable to identify its
antecedent in the coreferent-identification task. This pronoun was therefore eliminated from the coreferent-identification task in the main investigation.

Procedure

Subjects were tested individually. They were first presented a typewritten version of the Oates passage, and were asked to read it with the intent of understanding as much as possible. They were allowed to read the passage as many times as they wished, and to take as much time as they needed. When subjects indicated that they had finished reading the passage, they were given a blank sheet of paper and required to write, in their native languages, as much as they could remember from the passage.

This immediate-recall protocol, used to measure overall comprehension, allowed each subject to generate a reconstruction of the passage without the external influence of experimenter-generated questions. It thereby permitted an analysis of how each subject had processed information from the text. Subjects were told that their recall protocols would be translated for the examiner.

The reason for allowing subjects to write their recall protocols in their native languages was to avoid confounding an assessment of subjects’ ability to comprehend the language with their ability to produce it (Lee, 1986). The Spanish recalls were translated by the experimenter. Translators were paid to translate all other recalls. In order to ensure that the original meanings of the recall protocols had not been distorted by the translations, a sampling of the protocols was translated by a second translator. Reliability calculations, which were based on the total number of idea units recorded as present by each translator, revealed an agreement level of 92%.

After completion of this task, subjects were introduced to coreferential pronouns and were presented a double-spaced copy of the passage with the coreferential pronouns underlined (see Appendix B). They were required to write the antecedent of each of these pronouns above the pronoun.

The final portion of the examination consisted of an interview in which the subjects explained how they had interpreted the passage, especially focusing on meanings of the coreferential pronouns. At the end of each interview, inferential questions were posed. Also of interest was whether or not “compounding” effects—one wrong choice increasing the likelihood of further errors—might be observed. During this portion of the study, subjects, with the passage containing the underlined pronouns in front of them, were asked to explain the text, focusing especially on the pronoun
meanings. After completion of this explanation, the interviewer asked questions pertinent to textual comprehension and requested that unknown vocabulary be circled.

Scoring and Data Analysis

The scoring device for overall comprehension was constructed according to Meyer’s (1985) immediate recall protocol procedure. Application of this procedure for this study consisted of constructing a tree diagram representing the hierarchical ordering of propositions in the passage and transferring that hierarchical order to a scoring sheet in which all idea units of the passage are listed. The list of all idea units is presented in Appendix C. Because no significant difference was found between a scoring system that weighted idea units and the results of a simple count of units, the sum of items remembered was taken to represent each subject’s score. For the co-referent-identification task, raw scores were computed.

RESULTS

Summary statistics for overall comprehension and coreferent comprehension scores are shown in Table 1. A 3 X 2 multivariate analysis of variance procedure (MANOVA) was performed on these two dependent variables. No significant effects emerged due to language-proficiency level (Wilks’ $F[2, 26] = 2.68, p < .09$), language type (Wilks’ $F[2, 26] = 1.17, p < .33$), or the interaction between language-proficiency level and language type (Wilks’

<table>
<thead>
<tr>
<th>Proficiency level</th>
<th>Moderate ZAP languages</th>
<th>Maximal ZAP languages</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
</tr>
<tr>
<td>Comprehension</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prematriculation</td>
<td>62.86</td>
<td>14.19</td>
<td>49.17</td>
</tr>
<tr>
<td>Matriculation</td>
<td>61.14</td>
<td>18.19</td>
<td>54.45</td>
</tr>
<tr>
<td>Total</td>
<td>62.00</td>
<td>14.00</td>
<td>57.80</td>
</tr>
<tr>
<td>Coreferent identification</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prematriculation</td>
<td>23.39</td>
<td>5.28</td>
<td>21.83</td>
</tr>
<tr>
<td>Matriculation</td>
<td>25.86</td>
<td>3.63</td>
<td>25.09</td>
</tr>
<tr>
<td>Total</td>
<td>24.58</td>
<td>4.46</td>
<td>23.46</td>
</tr>
</tbody>
</table>

TABLE 1

Means and Standard Deviations of Comprehension and Coreferent-Identification Scores as a Function of Proficiency Level and L1 Anaphoric System
The null hypothesis, therefore, was confirmed for both comparisons (language level and language type).

That no significant differences were detected between proficiency levels suggests that the skills necessary for successful comprehension of natural text (i.e., successful integration of information) may not be sensitive to proficiency level, at least not above a threshold level of proficiency reached by these subjects. (For a discussion of a language-competence threshold necessary for the transfer of literacy skills, see Carson, Carrell, Silberstein, Kroll, & Kuehn in this issue of the TESOL Quarterly.) These findings are consistent with those of Bernhardt (1985) and Lee (1988), who found proficiency level to be an inadequate predictor of L2 reading comprehension performance. That the MANOVA procedure detected no significant differences attributable to L1 anaphoric system suggests that L2 readers from varying anaphoric backgrounds have, at this proficiency level, attained similar levels of proficiency with respect to overall reading comprehension and the comprehension of coreferential ties.

The Pearson product-moment correlation was .68 ($p < .000$) for the two dependent variables. This result indicates that there is a significant relationship between overall comprehension and the comprehension of coreferential ties for L2 readers. This positive correlation, together with the lack of main effect for L1 anaphoric system, indicates that linguistic processing skills do account for a portion—but only a portion—of competence in discourse comprehension for L2 readers.

The second phase of the study compared the performance of L2 readers with that of L1 readers. The comparison between overall comprehension and the ability to identify coreferential ties yielded a Pearson product-moment correlation of .55 ($p < .000$) for both L1 and L2 readers. This finding, along with the correlation of .68 for L2 readers, suggests that linguistic processing skills account for discourse comprehension less for L1 readers than for L2 readers. A discriminant function analysis was used as a follow-up procedure in order to explore possible reasons for these differences.

Discriminant function analysis is used to interpret multivariate data in order to explore group differences. This procedure was used to explore in what way each dependent variable (coreferent identification and overall comprehension) had contributed to the differences between groups. The procedure is based on an analysis of structure coefficients, which are correlations between the total variate and the constituent variables. The structure coefficients were .86 for the coreferent-identification task and .78 for the overall...
comprehension task, thus indicating that identifying coreferents contributed slightly more to the discrimination than did comprehension (propositions recalled). (Structure coefficients greater than .3 are considered to be significant.) The discriminant function achieved a Wilks’ lambda (1, 57) of .000.

The discriminant function analysis procedure correctly classified 72.88% of all subjects. It correctly classified 82.1% of the L1 readers and incorrectly classified 17.9% of them. Of the L2 readers, 64.5% were correctly classified as L2 readers, while 35.5% were incorrectly classified. The incorrectly classified L1 readers were those readers who achieved discriminant scores of less than .2383. Incorrectly classified L2 readers were those who achieved discriminant scores of more than zero. In other words, the lowest scoring L1 readers were predicted by the discriminant function as “belonging” to the L2 group, and the highest scoring L2 readers were predicted as “belonging” to the L1 group.

The above classification, then, permitted an analysis of four groups. Below are listed the four groups and the range of scores on each of the tasks:

1. L1 readers correctly classified as L1 readers (L1-L1)
   Overall comprehension: 55-146
   Coreferent identification: 28-29

2. L1 readers misclassified as L2 readers (L1-L2)
   Overall comprehension: 10-88
   Coreferent identification: 26-29

3. L2 readers misclassified as L1 readers (L2-L1)
   Overall comprehension: 60-88
   Coreferent identification: 27-29

4. L2 readers correctly classified as L2 readers (L2-L2)
   Overall comprehension: 16-70
   Coreferent identification: 16-29

The coreferent-identification task contributed slightly more to the discrimination than did the overall comprehension task. The L1-L1 readers included more idea units than the other groups and correctly identified at least 28 of the 29 coreferential ties. The L2-L1 readers also tended to include more idea units than the L2-L2 or L1-L2 groups, but, more important, approached the performance level of the L1-L1 readers on the coreferent-comprehension task. Of the five L1-L2 readers, all but one included very few (fewer than 56) idea units in their recall protocols. The remaining subject in this category included more idea units (88) in the recall protocol but correctly located only 26 of the coreferential ties. The L2-L2 readers
correctly identified 16-29 of the coreferential ties and included 16 to 70 idea units in their recall protocols. Although these readers included fewer idea units in their recall protocols than did the readers from the other groups, what distinguished them most from the other groups was the tendency to make more errors in identification of coreferential ties.

L1 Readers Performing as L1 Readers

Information gleaned from the recall protocols, coreferent-identification tasks, and interviews of L1-L1 readers indicated that these readers shared so many similarities that a prototype of their interpretation could be constructed. These readers understood the major lexical units and rhetorical relationships of the text. In the first paragraph, for example, all readers indicated comprehension of the following:

1. Eileen was walking the dog.
2. She was trying to stop the dog from tugging at the leash.
3. She thought of the dog as an extension of her father.
4. She was thinking of the similarities between her dog and her father.

These readers also indicated comprehension of rhetorical relationships such as the following:

1. Problem/solution
   - Problem: The dog tugging at its leash,
   - Solution: Eileen told the dog to “stop that.”
2. Comparison/alternative
   - Eileen thought of the dog not as an animal but as an extension of her father.

In other words, performance on the recall protocol and the coreferent-identification task indicated that these readers understood not only the concepts linked by the pronouns, but also which concepts were being linked.

A critical component of these readers’ performance, and one that distinguished them from many of the L2-L2 readers, was their comprehension of the anaphoric chain, her [Eileen’s] father . . . the dog’s master . . . the old man. All of these readers understood that these three descriptions referred to Eileen’s father. This knowledge apparently enabled these readers to integrate two major pieces of information: the descriptions of Eileen’s father in the first paragraph and, in the second paragraph, the fact that Marcey “belonged to” her father.
Inferences by L1-L1 readers were quite similar. All of these readers inferred, for example, that Eileen’s descriptions of her father and the dog were not flattering. Virtually all of the subjects, therefore, inferred that Eileen did not like the dog or her father very much. Two other points of agreement were that Marcey was the favored sister and that the mother appeared to be a weak person. When asked why they thought the mother might be pretending to be an invalid, subjects mentioned that it might be her way of escaping from either responsibilities of the household or having to deal with the father. All subjects in this group inferred that the household was an unhappy one and that Eileen was unhappy with the household.

L1-L1 readers varied in the extent to which they made inferences. Whereas some readers, when asked about the mother’s personality, for example, would infer that perhaps the mother had nothing to do with the dog because she too saw him as an extension of her husband, others would comment that they knew very little about the mother from only two paragraphs. Perhaps this difference accounts for the variation in coreferent identification for the pronoun referring to the inferred entity in the phrase, “it would be better if this dog died so that they could have some peace.”

In this group, 82% of the readers selected Eileen and her family as coreferents for this pronoun, but the remaining 18% selected Eileen and the squirrel as coreferents. That any readers in this group would select Eileen and the squirrel as coreferents is puzzling since 100% of these readers inferred that the family was unhappy. Perhaps these readers were reluctant to “take a chance” that the coreferent might be an entity beyond the text.

In sum, these readers integrated information from various parts of the text in order to build remarkably similar interpretations. Performances on the recall protocols, coreferent-identification tasks, and information obtained from the interviews indicated that all readers from this group understood the vocabulary and major rhetorical relationships in the text, and that they integrated this information in order to construct similar discourse models. Subjects seemed to differ only with respect to the extent to which they were willing to make inferences.

L1 Readers Performing as L2 Readers

What distinguished this group from the other L1 readers was either (a) less-detailed information on recall protocols or (b) poorer performance on the coreferent-identification task. A possible explanation for the less-detailed recall protocols might be that the
subjects were simply less able to integrate the information successfully, or that they chose to include only what they considered to be important or interesting. Of the five subjects in this group, only one incorrectly located a pronoun other than the inferential *they*. This subject selected *Marcey* and *Eileen* as coreferents of the pronouns *both* and *they* rather than *Marcey* and *the dog* in the lines:

The dog belonged in the house the way Eileen’s sister Marcey belonged: *Both* were possessions of the old man and could get away with anything. The fact was that *they* never really did anything at all. *They* belonged to the old man.

This subject’s interpretation differed from that of the L1-L1 readers: She perceived a shared experience between Eileen and Marcey, whereas L1-L1 readers saw a common experience (being possessed by the father) shared by Marcey and the dog, but not by Eileen. The negative connotations of a person “belonging to” or “being possessed by” another person (just as a dog belongs to a person) is absent from this reader’s interpretation. This absence affected the inferences made by this subject. Whereas the L1-L1 subjects noted Eileen’s disdain for her sister, this subject stated in the interview that she thought Eileen and Marcey had a good relationship: “[There is] some camaraderie between the two sisters because of them being able to get away with a lot of things.” This subject was the only L1 reader who did not comment that the household was unhappy. When asked during the interview how she would describe the household, she stated: “I would say a typical household. Any household has its rules and regulations.” Data from this reader provided information on how misinterpretation of one or two coreferential ties can affect overall comprehension. The failure to comprehend two coreferential ties affected perceptions of the relationship of the antecedents and, consequently, overall comprehension.

**L2 Readers Performing as L1 Readers**

The L2-L1 readers were selected by the discriminant function analysis procedure as performing comparably to the L1-L1 readers. These readers also demonstrated skill in using what knowledge they had to guess meanings of unknown items. Readers used knowledge of rhetorical relationships, for example, to guess—sometimes successfully—unknown lexical terms. For example, one subject’s awareness of the problem-solution relationship in the first paragraph allowed him to guess (correctly) that, because Eileen
tells the dog to stop doing something, he must be doing something annoying. The subject guessed (incorrectly) that the dog might be licking her hand: “She came out from the door with the dog with her, its tongue hanging, and it is licking her hand. She said, ‘stop it, you know, stop it’.”

One of the problems encountered by these readers, and also by the L2-L2 readers, was the comprehension of colloquial and/or figurative language. These readers knew, for example, the literal meanings of slow and belong, but did not understand their meanings as they were used in the passage. Another interesting result that emerged from this group of readers is that the comprehension of a coreferential tie does not guarantee that an L2 reader will draw the same inferences concerning the antecedent as would an L1 reader. One reader who understood all coreferential pronouns pertaining to the father . . . master . . . old man anaphoric relationship still did not understand that the protagonist harbored negative feelings toward her father, an inference shared by virtually all of the L1-L1 readers. In the recall protocol, this subject stated: “She never considered or treated the dog as an animal, rather she considered it an extension to her relationship with, or her love and respect for, her father.” During the interview this subject expressed confusion that Eileen wanted to get rid of the dog, even though she saw him as an extension of her father. The subject commented that it was a “strange family.” He found it difficult to accept that Eileen could express such negative feelings toward her father.

In sum, readers in this group demonstrated all of the attributes of skillful readers: the ability to perceive rhetorical relationship and to make intelligent guesses in order to guess unknown items, the use of linguistic knowledge in order to guess meanings, and an understanding of coreferential ties. The major obstacles seemed to be lack of familiarity with (a) figurative language and (b) elements of the target culture (e.g., that it is permissible to dislike one’s father).

L2 Readers Performing as L2 Readers

Scores of L2-L2 readers ranged from 16 to 70 on overall comprehension and 16 to 29 on the coreferent-identification task. Readers from this group made more pronoun errors than did those from the other groups. Therefore, pronoun errors made by this group of readers provided insights into which elements of discourse are most closely linked with comprehension of coreferential ties.

The majority of the coreferential-tie errors made by these readers consisted of the misinterpretation of (a) two pronoun clusters, that is, strings of pronouns, all of whose referents were misidentified,
and (b) a pronoun whose antecedent was not stated but rather had to be inferred from the text. The two pronoun clusters consisted of both . . . their . . . their . . . they . . . their from the first paragraph and both . . . they . . . they from the second paragraph:

She never thought of him as an animal, but as an extension of her father, whose dog he was; indeed, the dog resembled his master. Both had vague mottled skin, liverish and brown as if camouflaged, and their eyes were watery with alertness. It seemed to Eileen that their ears, though of different shapes and colors, had in common an unclear, intangible quality of intensity—they both heard everything, heard whispers not meant for their ears and words not spoken aloud, heard even the echoes of words that should have faded away.

The dog belonged in the house the way Eileen’s sister Marcey belonged: both were possessions of the old man and could get away with anything. The fact was that they never really did anything at all. They belonged to the old man.

Readers who misunderstood the link between Eileen’s father and the dog’s master also misunderstood the subsequent pronouns both, their, their, and they, which were used to refer to the dog’s master and the dog. Therefore, these readers did not understand that the dog and master were being compared. As a consequence, they were not aware of this subtly derogatory description of the father. Of the L2-L2 subjects, 40% misunderstood the first cluster, and 85% misunderstood the second cluster. Although only a minority of the L2-L2 subjects (40%) misunderstood the first cluster, 88% of those who misunderstood it also misunderstood the pronoun cluster in the second paragraph. Stated another way, 60% of the readers understood the first cluster, but became confused somewhere between the end of the first paragraph and the beginning of the second pronoun cluster in the second paragraph. A strong majority (88%) of those who misunderstood the first pronoun cluster never did “recover.” It seems, then, that comprehension of the second cluster depended, to a large extent, on comprehension of the first cluster. Understanding that the old man referred to the same person who was described in the first paragraph was necessary in order to integrate information from both paragraphs. More specifically, knowing that Eileen’s father was the dog’s master and that Eileen identified him with the dog apparently helped readers to understand the analogous relationship between the dog and Marcey in the second paragraph. The readers who did not understand that father, master, and the old man referred to the same person not only were unable to integrate information pertaining to this character from all parts of the passage, but were forced to construct
interpretations that would accommodate additional characters. The following are recall protocols from L2-L2 readers who shared this problem.

The first protocol illustrates the situation of a reader who misinterpreted the author’s usage of master as a specific person from the text, and then was unable to locate the appropriate coreferents for the pronouns of the pronoun cluster.

The woman does not think that a dog can be an interesting animal. But her father told her that a dog could be loyal to its master. The dog has a great sense of hearing. Even to whispers, the dog could overhear the woman walking along the street taking her parcel.

Without an understanding of the descriptions of Eileen’s father in the first paragraph, the reader whose recall protocol follows was unable to see that the old man was the same person. She stated: “But the dog belonged to her sister, Marcey. And also the dog belonged to an old man who stayed with Marcey. Marcey lived with her mother.”

Another reader wrote the following:

The main character, Eileen, comes face to face with a strange situation in which she discovers that strange animals came to be extensions of people, which in turn were under the control of a Master. Specifically, she encounters the case of her father who, in the form of an animal with colored ears and alert eyes, was pestering her. She was very surprised because she couldn’t imagine that that animal could be an extension of her father.

This reader, who named Marcey and Eileen’s mother as coreferents of the second pronoun cluster, did not mention the old man in the recall protocol. When asked during the interview who the old man was, he replied: “At the beginning I thought it was the master, but it also could be the old father. I’m confused—whether it’s the father or the master.”

Some readers interpreted the pronoun cluster of the first paragraph as referring to Eileen and the dog, possibly because they assumed that Eileen was the dog’s master. Below are two excerpts from protocols of readers who made such an interpretation and, subsequently, created another old man:

After that, she went to the long side of the street. There was an elderly man who lived on that street. He owned the dog.

It happened that the houses along the street had dogs, too. An old man, who felt it unnecessary to keep his dog, was bothered because he had to feed it with bones.

Some readers who understood the pronoun cluster of the first paragraph were still unable to see the link between Eileen’s father
and the old man. One such reader incorporated old into his interpretation by deciding that the house where Eileen and her sister grew up now belonged to two old persons:

She stopped at a house where she and her sister grew up. Now the house belonged to two old persons. She made the dog stop at the door of the house. The two old persons do nothing to the dog. They were very old already. People could enter the house to take anything they wanted (or to take as many things as they could).

The pronoun-cluster errors described above suggest that a misunderstanding of coreferential ties can be a manifestation of a misinterpretation of antecedents. In these data, readers who misunderstood pronominal referents had misunderstood the descriptive expressions to which the pronouns referred.

A third problematic coreferential tie for L2 readers was the pronoun they in the sentence: “The dog wanted so badly to tear flesh with his jaws that Eileen stared at his flat, brutal skull and thought clearly that it would be better if this dog died so that they could have some peace.” This pronoun was particularly dependent upon inferential reasoning; the selection of an appropriate antecedent has to be made from outside the text. In other words, anaphoric resolution depended upon an awareness of topic. Prince (1981) classifies discourse entities into three general categories: new, evoked, and inferable. She terms inferrable those entities whose antecedents must be inferred from the text, claiming this category is the “most complex” (p. 236) of the three. It is therefore not surprising that the majority of the L2-L2 readers misunderstood the pronoun they, discussed above.

In scoring, the selection of Eileen and her family, Eileen and her mother, and Eileen and Marcey were judged acceptable coreferents, because the protagonist had expressed negative feelings toward the dog and her father. The selection of such antecedents as Eileen and the dog, Eileen and the squirrel were judged as unacceptable responses. L2 readers who selected the latter type of antecedent may have done so because L2 readers are traditionally taught to expect a pronoun’s antecedent to be the nearest preceding noun phrase, and the choice of the nearest preceding noun phrase is most likely to be correct if one does not understand the passage.

Bosch (1983) points out that “we could hardly say any less or presuppose any less about something in referring to it than in the case where we use a pronoun” (p. 203). He goes on to say:

The intended referent for an anaphoric pronoun thus must already be the most salient object in the domain of discourse at the moment of the utterance, more salient at least than any other object that could possibly
be referred to by the same form. And the most salient object in any
domain of discourse is the object the discourse is about at the relevant
moment.

Apparently, the L2 readers did not perceive the discourse to be
“about” Eileen’s family in this particular instance.

DISCUSSION

The L1 interviews and task performance provided an invaluable
comparison to those of L2 readers. The L1 data provided informa-
tion concerning how a “typical” L1 reader integrates various types
of information (e.g., culture, syntax, vocabulary) in order to
construct a discourse model for a particular text. A comparison of
L2 data with that of L1-L1 readers revealed that comprehension
problems regarding anaphoric chains are two-pronged. First, L2
readers encounter difficulties when they are unfamiliar with the
descriptive expression used as the antecedent of a coreferential
pronoun. Second, lack of comprehension of these expressions may
be indicative of a lack of knowledge of the target culture. In this
particular case, L2 readers who misunderstood the usage of the old
man may have also misunderstood that it is permissible in United
States culture to harbor negative feelings toward one’s father.
Results from this study suggest, then, that comprehension problems
with respect to coreferential ties may stem from (a) lack of
familiarity with a descriptor or (b) lack of knowledge of the cultural
concept that a particular descriptor conveys.

The cultural component, however, surely represents only a partial
explanation of the comprehension problems described above. Examin-
ation of the L1-L1 readers’ interviews and task performance
indicates that such factors as background knowledge, vocabulary,
and knowledge of rhetorical relationships interact to aid the reader in
constructing an interpretation of a passage. Whereas an L1 reader
may rely on the interaction of all of these components, an L2 reader
may have access to only one or two of them in a particular instance.
The success or failure of resolving anaphoric relationships depends,
in large measure, on the successful interaction of all of these
components of the reading process. Further research is needed in
which the interaction of two or more of these components for L2
readers is compared with that of L1 readers.

PEDAGOGICAL IMPLICATIONS

The pedagogical implications derived from these studies can
perhaps best be summarized by the comment of one L2 subject during the interview: “I know what all of these words mean individually, but I still don’t understand the passage.” His comment, which supports the statistical findings of this study, indicates that proficiency level does not provide the total picture. Students need help both in locating pronominal referents and in developing strategies for interpreting these.

Teachers can also provide needed conceptual/cultural information in prereading activities. Instructors can search out possible misinterpretation problems regarding different types of concept designators, particularly figurative and colloquial expressions. Prereading can begin, when possible, with discussion of universally known concept designators (e.g., her father), and proceed to culturally known (e.g., the old man) expressions and their connotations. After students read, teachers can supplement cultural discussions with additional reading materials—for example, short newspaper and magazine features, advertisements, and cartoons—that convey the cultural values, and, if relevant, the humor and/or irony associated with a concept. Finally, students can be given the opportunity to read more works of the same genre or by the same author so that they can apply their newly gained background knowledge to the interpretation of similar texts.

Not all problems with coreferential ties, of course, can be prevented with prereading instruction. Nor is all error making negative. Errors can offer opportunities for explicit work on comprehending strategies, and can provide diagnostic information for teachers. When errors are made, teachers can use these errors to detect other comprehension problems in the text. Most important, students should be taught that confusion in interpreting a pronoun may indicate a misunderstanding of another entity in the passage. Such an emphasis would lead L2 readers to monitor their own comprehension and ultimately to become more actively engaged in integrating all available information for successful comprehension of a text.

THE AUTHOR

Marjorie C. Demel is Assistant Professor of Spanish at Denison University. She has conducted research on second language reading comprehension of French, Spanish, English, and German.

READING COMPREHENSION AND COREFERENTIAL TIES

285
REFERENCES


APPENDIX A

Overall Reading Comprehension Immediate Recall Protocol

The following is an introductory passage to a short story, “Demons,” by Joyce Carol Oates. Please read the passage with the intent of understanding as much of it as possible. Read it as many times as you wish and take as much time as you need. After you have finished reading it, hand it to the examiner. You will then be asked to write everything you can remember from it.

DEMONS*

She came out the side door, with the dog tugging at his leash. The worn leather grip jerked her hand and she said in a hushed, embarrassed voice: “Stop that, you know better! Stop that!” The dog was vain and partly bald. She never thought of him as an animal, but as an extension of her father, whose dog he was; indeed, the dog resembled his master. Both had vague mottled skin, liverish and brown as if camouflaged, and their eyes were watery with alertness. It seemed to Eileen that their ears, though of different shapes and colors, had in common an unclear, intangible quality of intensity— they both heard everything, heard whispers not meant for their ears and words not spoken aloud, heard even the echoes of words that should have faded away.

She took the long side walk to the street, holding back on the leash. The dog panted and yipped in a high womanish falsetto; there was a squirrel nearby. The dog wanted so badly to tear flesh with his jaws that Eileen stared at his flat, brutal skull and thought clearly that it would be better if this dog died so that they could have some peace. The dog belonged in the house the way Eileen’s sister Marcey belonged: both were possessions of the old man and could get away with anything. The fact was that they never really did anything at all. They belonged to the old man. Eileen’s mother was an invalid, or claimed to be, and she had nothing to do with the dog, which was not allowed above the landing—nor was the dog allowed in the parlor, etc., etc.—these were the rules of the house that Eileen had grown up with and had supposed to be the laws of the city itself, each house on each street encumbered with bawling vicious dogs and fathers and mothers and sisters, one sister good because she was “slow” and the other—Eileen herself—questionable because she was not slow.

APPENDIX B
Coreferent-Identification Task

Pronouns frequently refer to items or ideas, called “coreferents,” found elsewhere in the text. In the sentence pair (a) below, for example, “she” refers to “Mary.” In the sentence pair (b), “her” refers to “Mary.”

a. Mary is a student at Ohio State. She lives in Jones Tower,

b. Mary studies at Ohio State. Her brother studies there, too.

The following sentence pairs contain more examples of pronouns and their coreferents. Each pronoun’s coreferent has been written above the pronoun,

*(the food in the cafeteria)*

c. Jim loves the food in the cafeteria, but Mike hates it.

*(living in the dorm)*

d. Jim loves living in the dorm, but Mike hates it.

*(Maria)*

e. Because she will be studying in the United States next year, Maria needs to learn English.

Practice example:
Find the coreferent for the underlined pronoun in the sentence pair below. Write the pronoun’s coreferent above the pronoun.

f. John has to get up early to go to school.
   
   His first class meets at 8:00.

The pronouns in the passage on the next page have been underlined. Please write above each pronoun the item or idea that pronoun refers to.

Demons

She came out the side door, with the dog tugging at his leash. The worn leather grip jerked her hand and she said in a hushed, embarrassed voice: “Stop that, you know better! Stop that!” The dog was vain and partly bald. She never thought of him as an animal, but as an extension of her father, whose dog he was; indeed, the dog resembled his master. Both had vague mottled skin, liverish and brown as if camouflaged, and their eyes were watery with alertness. It seemed to Eileen that their ears, though of different shapes and colors, had in common an unclear, intangible quality of intensity— they both heard everything, heard whispers not meant for their ears and words not spoken aloud, heard even the echoes of words that should have faded away.
She took the long sidewalk to the street, holding back on the leash. The dog panted and yipped in a high womanish falsetto; there was a squirrel nearby. The dog wanted so badly to tear flesh with his jaws that Eileen stared at his flat, brutal skull and thought clearly that it would be better if this dog died so that they could have some peace. The dog belonged in the house the way Eileen’s sister Marcey belonged: both were possessions of the old man and could get away with anything. The fact was that they never really did anything at all. They belonged to the old man, Eileen’s mother was an invalid, or claimed to be, and she had nothing to do with the dog, which was not allowed above the landing—nor was the dog allowed in the parlor etc., etc.—these were the rules of the house that Eileen had grown up with and had supposed to be the laws of the city itself, each house on each street encumbered with balding vicious dogs and fathers and mothers and sisters, one sister good because she was “slow” and the other—Eileen herself—questionable because she was not slow.

APPENDIX C
Scoring Instrument for Recall Protocols
List of Content and Relationship Units

DESCRIPTIONS OF ACTIONS
CAME OUT SIDE DOOR, TOOK SIDEWALK
agent SHE (EILEEN)
range STREET
description: manner TUGGING AT LEASH
agent DOG
description: attribution VAIN
description: attribution PARTLY BALD
problem/solution: problem JERKED HER HAND
agent GRIP
description: attribution LEATHER
description: attribution WORN

Note. The process of developing a tree diagram representing the hierarchical ordering of propositions in the passage revealed two categories of propositions: descriptions of actions and descriptions of thoughts. Uppercase entries above represent relationship units; lowercase entries indicate context units. Subjects received one point for each rhetorical relationship and each lexical item recalled.

READING COMPREHENSION AND COREFERENTIAL TIES
problem/solution: solution
SAID “STOP THAT!” ETC.
description: manner
HUSHED
EMBARRASSED
description: manner
HOLDING BACK ON LEASH
causation: covariance. consequent
PANTED
YIPPED
description: attribution
FALSETTO
description: attribution
HIGH WOMANISH
causation: covariance, antecedent
THERE WAS SQUIRREL
range
NEARBY
causation: covariance, antecedent
WANTED TO TEAR FLESH
description: manner
SO BADLY
WITH JAWS
causation: covariance, consequent
STARED AT SKULL
agent
EILEEN
description: attribution
FLAT
BRUTAL

DESCRIPTIONS OF THOUGHTS

THOUGHT
agent
EILEEN
comparison: alternative
NOT ANIMAL
patient
DOG
comparison: alternative
EXTENSION OF FATHER
description: attribution
WHOSE DOG HE WAS (BELONGED TO FATHER)
description: representative identification INDEED
RESEMBLED FATHER
patient
DOG
collection
HAD SKIN
patient
BOTH
description: attribution
VAGUE MOTTLED
BROWN AND LIVERISH
description: attribution
AS IF CAMOUFLAGED

290  TESOL QUARTERLY
HAD EYES
  description: attribution
  WATERY WITH ALERTNESS
  (HAD) EARS
  comparison: adversative THOUGH
  DIFFERENT
  description: attribution
  SHAPES
  COLORS
  comparison: alternative
  (HAD) IN COMMON
  patient
  INTENSITY
  description: attribution
  UNCLEAR
  INTANGIBLE
  HEARD EVERYTHING
  description: representative identification
  WHISPERS NOT MEANT FOR THEIR EARS
  WORDS NOT SPOKEN
  ECHOES OF WORDS
  description: attribution
  SHOULD HAVE FADED AWAY
causation: covariance, antecedent
  BETTER IF DOG DIED
causation: covariance, consequent
  COULD HAVE PEACE
  comparison: analogy
  BELONGED IN HOUSE
  patient
  DOG
  comparison: analogy THE WAY
  BELONGED IN HOUSE
  patient
  MARCEY
  description: equivalent
  EILEEN’S SISTER
causation: covariance, antecedent
  WERE POSSESSIONS OF OLD MAN
causation: covariance, consequent
  comparison: alternative
  COULD GET AWAY WITH ANYTHING
  comparison: alternative
  NEVER DID ANYTHING AT ALL
causation: covariance, antecedent
  comparison: akernative
  WAS INVALID
comparison: alternative OR
  CLAIMED TO BE INVALID
  patient and/or agent
  EILEEN’S MOTHER

READING COMPREHENSION AND COREFERENTIAL TIES
causation: covariance, consequent
HAD NOTHING TO DO WITH
patient
DOG
description: attribution
NOT ALLOWED ABOVE LANDING
NOT ALLOWED IN PARLOR
ETC., ETC.
description: equivalent
WERE RULES OF HOUSE
description: equivalent
HAD GROWN UP WITH
HAD SUPPOSED TO BE LAWS OF CITY ITSELF
ENCUMBERED WITH
patient
EACH HOUSE
description: setting location
EACH STREET
description: attribution
DOGS
description: attribution
BALDING
VICIOUS
FATHERS
MOTHERS
SISTERS
comparison: adversative
causation: covariance, consequent
ONE GOOD
causation: covariance, antecedent
ONE SLOW
comparison: adversative
causation: covariance, consequent
ONE QUESTIONABLE
causation: covariance, antecedent
NOT SLOW
Recent Publications on Statistics, Language Testing and Quantitative Research Methods: II

Fundamental Considerations in Language Testing

Assessing Achievement in the Learner-Centred Curriculum

Testing for Language Teachers

In the first part of this article (TESOL Quarterly, Vol. 23, No. 1), I reviewed four books, those by Brown (1988); Butler (1985); Henning (1987); and Woods, Fletcher, and Hughes (1986). I stressed the importance of the ability of applied linguists and language teachers to read the literature of their field evaluatively, without being intimidated by its quantitative aspects. I also stressed the importance for language teachers of reaching sound testing decisions that respond to specific educational contexts, either with existing tests or with teacher-developed exams. In this second section, I review three more volumes on testing.

Bachman’s large work, Fundamental Considerations in Language Testing, begins with the statement that “this is not a ‘nuts and bolts’ text on how to write language tests” (p. 1); the reader will find this true. What, then, is it? Bachman describes it as a “discussion of fundamental issues that must be addressed at the start of any language testing effort” and an exploration of “some of the problems raised by [the fact that] language is both the instrument and the object of measurement” (p. 1). Bachman identifies three fundamental considerations: the contexts determining the uses of language tests; the nature of language abilities; and the nature of
measurement, focusing throughout on the measurement of language proficiency.

I begin with the last chapter because this is, for me, the most interesting and satisfying. It is also the chapter where Bachman speaks directly to his audience and identifies, as I do not think he does until that point, precisely who his audience is. Here Bachman discusses “authentic language” tests (p. 300), and develops a distinction between what he calls “real-life” (RL) and “interac-tional/ability” (IA) tests (pp. 301-302), which is immediately convincing and useful. Around this contrast Bachman builds a wide-ranging and valuable discussion of problems and directions in language testing, located firmly within the discourse community of language-testing researchers. The persistent problems and future directions he discusses, and the terms in which he considers them, are not those of teachers who must develop or select language tests for their own students. This chapter clarifies for me the uneasy feeling I had throughout the book, that of needing to identify the intended audience.

There is much of value in Bachman’s book for practicing teachers and students in TESOL courses. I suspect, however, that teachers, reading alone, will find the book hard going indeed, and that master’s degree students will be grateful for a well-prepared instructor to guide them through the text, in particular helping them identify concrete cases, problems, and data sets that can bring into perspective the theoretical issues discussed by Bachman. This need for supplementary example is reminiscent of Henning’s volume discussed in the first part of this review; however, the book is unlike Henning’s in most regards. Bachman allows himself full range to consider, at length, the fundamental issues that concern him, and is unabashedly willing to dismiss issues of less interest with but brief comment. This is a personal book, making no attempt at complete coverage, of everything a graduate student might want to know about language testing. While this approach makes the book interesting for the committed reader, it may present a problem for anyone designing a one-semester course in language testing who aims to cover specific ground in the way recommended in the first part of this review.

Chapter 1 is not, for me, the book’s strongest. The brief sections on second language acquisition research, language ability, and measurement theory leave me wondering what central point I should be carrying away to apply to my reading later in the book. It seems that several concerns are raised, only, to be left hanging, while we wonder what position Bachman holds, or what he assumes us to have inferred from the discussion. A figure borrowed from
Chapter 6 attempts, it seems, to link the several short discussions of areas that especially interest Bachman to some sort of philosophical base or overall plan, but the reader without a solid grounding in these areas will find the discussion obscure. Placing the overview of the book earlier might have helped to clarify this chapter. In fact, the book is cohesive, but to learn this it is necessary to read all of it. This is not, in my view, a book that is made to dip into, in the way that many of us have become accustomed to doing with sourcebooks for testing courses.

The chapters that deal with measurement are excellent, although the incorporation of work with data into the text, or the provision of concrete examples for the theoretical arguments and explanations Bachman is presenting, would help a language-testing initiate make better sense of the discussion. But, as I have indicated above, I do not believe this is a book for the uninitiated. Chapters 2 and 3 cover more or less typical ground for a book on language testing, but, again, Bachman’s individual concerns are evident throughout. The sections on “Steps in Measurement” in Chapter 2 and “Research Uses of Tests” in Chapter 3 are especially welcome. Bachman does not, however, attempt to introduce or teach any actual statistical techniques for working with quantitative data. The book holds true to the promise of its title in focusing on considerations, not applications.

Bachman takes (in Chapter 6) a wider view of reliability and threats to it than do many more traditional books on language testing; he provides a clear and quite detailed explanation of classical measurement theory, generalizability theory, and item response theory, while providing statistical formulae for concepts under consideration. In discussing validity (in Chapter 7), he follows recent practice in stressing construct validation. While he discusses correlational, factor analytic, and multitrait multimethod approaches to construct validation, he does not provide either the statistical formulae or any concrete examples of these.

In Chapters 4 and 5, Bachman introduces and develops in detail his framework for communicative language ability and “test method facets” (p. 115), which has already been influential in language testing through dissemination and discussion at testing meetings, and in its application in the Cambridge-TOEFL Comparability Study (Bachman, Davidson, & Foulkes, 1990). In large part the framework is presented descriptively and in language that appears to imply an a priori reality for the model. In fact, it could well provide researchers in language testing with researchable issues beyond the end of the century, and indeed Bachman himself has suggested that something of the sort should happen.
Bachman and Clark (1987). Bachman's framework is theoretically interesting in a number of ways, but my own feeling is that it would have received the attention it deserves more easily if presented in a separate forum. Its presence in this more broadly construed text confirms for me that this is the book Bachman wanted to write, and that it must be read on that basis.

Brindley's *Assessing Achievement in the Learner-Centred Curriculum* was prepared as a report to the Australian Adult Migrant Education Program (AMEP), an Australian government program serving over 130,000 learners, primarily adult ESL learners, each year. Brindley makes no claims that this book could or should be used as a textbook for language-testing courses, and thus my review of it here may seem unfair. Nevertheless, Brindley has provided us with a powerfully informative example of language testing research and language tests as pedagogical problem solving in action. In the report, Brindley investigates the kinds of information on learner achievement that are needed at all levels of the AMEP, and the ways that fine-grained assessment of learners' language gains over courses of instruction can be carried out. As part of this, he looks at assessment methods in other contexts and considers their applicability or adaptability to the needs of the AMEP.

Brindley provides a fascinating look at teachers' responses to a survey of their views on the importance of assessment in their work. He found that placement into classes figured highest, with diagnostic information for use in course planning second. Looking at assessment methods favored by teachers, he found that informal methods ranked highest (with observation followed by recycling of work ranked most highly) and informal discussions with learners second; standardized testing ranked lowest. Teachers were very interested in learner self-assessment, although most of them felt underequipped in applying this. Interestingly, self-assessment is scarcely dealt with at all in the books I have reviewed here and in Part 1, although it is becoming an increasingly important movement, particularly in Europe and Australia.

Self-assessment is just one form of assessment found appropriate within the learner-centered educational philosophy held by the AMEP. The book contains the best discussion of learner profiling I have encountered, along with reference, although limited, to the folio of learner work that parallels the developments in portfolio assessment currently enjoying popularity in U.S. composition assessment. While Brindley's discussion of accountability may be rather obscure to those of us outside AMEP, it does demonstrate that while some of our language testing problems can be solved
with reference to a wider perspective of literature, research, and experience, others must perforce be solved more introspectively.

The bulk of Brindley’s book is devoted to a detailed description of, and argument for, various criterion-referenced achievement assessment instruments, a wealth of which are presented within the middle chapters of the book, in a way that is both fully contextualized and firmly based theoretically. Contained in this book is enough project material for several language testing classes to explore in practice many of the issues raised in a more abstract way in Bachman’s book. This volume should prove useful supplementary reading for any language testing course, (Teachers seeking another book to fill this supplementary niche might also examine Weir’s Communicative Language Testing [1988].)

Hughes’ Testing for Language Teachers starts with a clear specification of its audience and purpose: The book is intended “to help language teachers write better tests” (p. ix). The text is consistent with this purpose in its tone, level of complexity, and limits on its assumptions of prior knowledge and experience. I note, however, that there is almost no reference to statistics throughout the book, relegating statistical information that does appear to an appendix. I believe that teachers can handle more statistics than Hughes appears to give them credit for, and that indeed they must. How can teachers design effective tests if they have no tools with which to evaluate their efforts? With almost every section of this book, specialists in language testing will want more; but for short in-service courses and for practitioners using the book in their home contexts to provide guidance as they actually solve their testing problems, what Hughes offers is probably enough, except in limits he places on statistical tools. I suspect, however, that even those using the book in short courses will need to supplement the book with some manipulation of quantitative data.

Hughes’ explanations of basic concepts are brief but clear and pertinent. In Chapter 3, I note in particular a helpful discussion of the differences between kinds of language tests, with welcome attention to achievement versus proficiency, and norm-referenced versus criterion-referenced tests, and a welcome dismissal of the objective/subjective testing myth in a few lines. The further readings are quite well chosen to offer more detail in each area but, in most cases, without becoming too technical. But the very limited extent of the discussions does cause me some concerns: As an example, Chapter 4 deals with validity. A sound grasp of validity is especially important for a practitioner audience, both because of its classroom impact and because many teachers will have previously
studied language testing, if they have studied it at all, in an older paradigm where validity was underemphasized relative to reliability. I would have liked to see more attention to the argument for an emphasis on validity in testing and test design, and some discussion of the operational implications of such an emphasis. Similarly, in Chapter 5 on reliability, the techniques discussed are fine, but given the stress placed in the book on teacher-made tests and the recurring advice (with which I agree) to avoid multiple-choice testing, I would have expected a different balance. It seems curious in this chapter to place the focus of concern away from the kinds of criterion-referenced testing techniques Hughes has declared himself to favor.

For me, the weakest chapter of the book is perhaps the most important one: Chapter 7 on “Stages of Test Construction.” There is so little detail in the main text that too much weight may be given to the examples, neither of which does much to illustrate how the principles Hughes has expounded can be put into operation. Further, no mention is made of a stage after the test has been administered: Practitioners need to learn how to make use of test results, and how to evaluate those results to plan for improved testing in the future.

The remaining chapters on test techniques and the testing of specific language skills contain useful advice and examples. The final chapter, on test administration, will be a real help to teachers attempting serious testing in their classes or schools for the first time.

Despite the limitations I have described, this is an excellent small book, probably the first serious British contender as a replacement for the classic Writing English Language Tests (Heaton, 1975), a book that is still used widely for courses in language testing for teachers. Teachers looking at the Hughes book should probably also look at Heaton, and at Testing Language Ability in the Classroom (Cohen, 1980), and Techniques in Testing (Madsen, 1983).

What pleases me most about these three books is their shared emphasis on the ethical imperative of language testing. In teaching a course in language testing, at any level, it is critical to emphasize test uses and the power of test-based decisions. Teachers and applied linguists—anyone who is likely to read testing research literature and draw pedagogical or bureaucratic conclusions from it, or to evaluate existing tests, or to develop new tests—must be provided with ways to consider what tests mean, in terms of the methods used, the abilities elicited, the societal and individual consequences that ensue, and so on. Bachman’s book provides a
strong theoretical foundation for doing this; in Chapter 7 he provides a highly relevant overview of ethical considerations in language testing. Bachman does not directly engage the specific issues, instead he provides a summary overview, rightly indicating that this is an area where language testing needs to focus much more effort. Brindley does this implicitly at every point in his book, and explicitly at the many points where he discusses accountability and collaborative assessment. The AMEP assessments serve as an example to testers and teachers alike, of what can be achieved by careful thought to what learners, teachers, and other involved parties need from a language assessment, and how testing experts can work with them to achieve it. The Hughes book begins with a discussion of backwash (also known as washback), or the effects of testing, beneficial and otherwise, on curriculum. Hughes argues that good tests are a powerful tool for curricular improvement. He also argues that teachers, students, and institutions (schools/colleges) need tests, and thus teachers must take responsibility for better testing. Hughes reiterates these views throughout his book and, despite its limitations, the book goes a pleasing distance in attempting to illustrate how this perspective can be turned into good testing practice.

Surveying these three books and the four previously reviewed, I realize that comparing books in this area is difficult because each is different in both approach and goals. In order to begin to compare textbooks in this area, questions about audience become critical. Lazaraton, Riggenbach, and Ediger (1987) found that 47% of their respondents, all of whom had established some researcher credentials in TESOL, had taken only one “tools” course. It seems most likely, then, that readers of these books also will have taken, or will be taking, only one course in language testing, or statistics, or research design. I have not located, among those I have reviewed, the perfect coursebook for a single-semester course in language testing. Another book eagerly awaited by language testers is Hatch and Lazaraton’s The Research Manual: Design and Statistics for Applied Linguistics, expected soon. This book promises to fill a major gap, but again will not be all things to each of us. The perfect coursebook, for me, would have Henning’s or Hughes’ coverage of key practical concepts, Brindley’s learner focus and wealth of examples, Bachman’s level of theoretical analysis of issues facing language testers, Woods et al’s (or Hatch & Lazaraton’s) level of statistical explanatory power, and Brown’s attention to interpretation of the research literature. Is this possible in one volume? And if it were, could it be taught or learned in a single semester? Realistically, teachers and teacher-trainers must make difficult
decisions about inclusion and exclusion, about ethos and values, about the kinds of graduates one wants a program to produce, and must select from the increasing array of materials in statistics, language testing, and quantitative research methods in order to match our requirements. The choices are better than they have been at any time in the past.

ACKNOWLEDGMENTS

I am grateful to Andrew Cohen, Dan Douglas, Larry Selinker, and Bernard Spolsky for comments that have assisted me in my review. All opinions, of course, remain my own.

REFERENCES


LIZ HAMP-LYONS

University of Michigan
University of Colorado at Denver (as of fall 1990)
Breaking Rules: Generating and Exploring Alternatives in Language Teaching


Breaking Rules appears at a time when people engaged in foreign language acquisition and teaching have become accustomed to experimenting with an enormous variety of innovative ideas, approaches, and methods. Taking into account the results of recent research, Fanselow’s up-to-date book not only summarizes and reveals insights into this extensive and diversified field, but also offers numerous practical applications for classrooms and other instructional settings.

The content of the book is certainly in keeping with the subtitle: Generating and Exploring Alternatives in Language Teaching. There is abundant advice on teaching in specific situations. There are also numerous examples of solutions to various instructional dilemmas, from the presentation of grammar and vocabulary to the negotiation of sociocultural barriers. All of these aspects of language teaching are discussed or, in Fanselow’s terms, generated and explored at great length, and often in minute detail. The author generally seems eager to keep his promises regarding clarity, manageability, and abundance of information: This last intent, however, leads occasionally to an exhaustiveness that appears to override the former concerns. The small print adds to the probability that readers will not find this manual entirely easy to follow.

While Breaking Rules does not advocate revolution in classroom procedure, the book is obviously intended to provide various means of breaking the monotony of routine classroom practice. According to Fanselow, breaking rules is a process that inevitably comes full circle, establishing its own rules and routines—these to be broken in turn. This book offers as its premise that any changes to be made, or alternatives explored, must be preceded by a clear understanding of the way things currently work. In order to provide this understanding, Fanselow offers a systematic instrument for observing and coding units of communication. This coding scheme called Foci for Observing Communications Used in Settings (FOCUS), is based on Bellack’s (1966) concept of the structure of classroom discourse, and is comparable to the Sinclair and Coulthard framework (1975). Teachers are invited to participate actively while reading, recording lessons, making journal entries, and engaging in systematic observation of lessons.

The center of this book is language use. Twenty-four “Episodes”
form the complete work, each illustrating a particular dimension of
the FOCUS scheme for observation of communication; illustrations
include examples of instructional and nonpedagogic discourse. All
of the episodes have essentially the same structure: First, a short
summary introduces new ideas (in most cases accompanied by an
outline); the episode then centers on excerpts of recorded and
transcribed audiovisual data, including paralinguistic and nonverbal
features (the categories and codes used in the episodes are listed on
the inside pages of the book cover).

Although *Breaking Rules* provides invaluable information on
teaching languages and offers ample reference to the author’s
sources, the text and format are highly demanding and complex,
perhaps rendering it an unsuitable manual for teacher trainees.
First, the dense concentration of information may appear daunting
to the neophyte. Second, the FOCUS scheme, while perhaps
appropriate for classroom researchers or for experienced teachers,
is of such complexity and detail that it is a cumbersome instrument
for teacher training. One aspect of the book that helps to
compensate for its length and textual density is the marking of
nonessential sections. The reader who does not have ample time for
intensive exploration is advised to skip those sections.

In summary, this book contributes little that is new to discussions
of teaching methods and approaches. For its methodological basis,
it depends upon currently accepted principles of communicative
language teaching, well explained elsewhere (e.g., Kramsch, 1988,
1985, 1981; Savignon, 1983; Rivers, 1987). What is unique about
*Breaking Rules* is its emphasis upon analysis of classroom discourse:
the inclusion of transcripts of naturalistic audiovisual data may
provide teachers fresh insights into the world of the classroom and
beyond.

REFERENCES

College Press.
Kramsch, C. J. (1988). To group or not to group? Le parler, le faire et le
dire en travail de groupe. *Langues Modernes, 6*, 16-23.
Rivers, W. M. (1987). Interaction as the key to teaching language for com-
munication. In W. M. Rivers (Ed.), *Interactive Language Teaching*
(pp. 3-16). Cambridge: Cambridge I.University Press.


RUTH SPYCHER

*Massachusetts Institute of Technology*
New Books from TESOL Publications!

Coherence in Writing: Research and Pedagogical Perspectives
Ulla Connor and Ann M. Johns, editors

This 12-chapter collection offers the most comprehensive and varied treatment of coherence available. The authors interpret the elusive term “coherence” for the ESL/EFL teacher, teacher trainer, and researcher in L1 and L2. The text covers both the processes of making and understanding coherence and the manifestations of coherence in written products. Piloted discussion questions and activities follow each chapter.

Students and Teachers Writing Together: Perspectives on Journal Writing
Joy Kreeft Peyton, editor

This book is about writing—how students use writing to explore and develop their own thoughts and interests, and how students and teachers can work together to understand each other and the world through their writing. The authors examine journal writing as a powerful stimulus for curriculum and teacher change; as a forum for discussing the meaning and form of literary texts and the journal texts themselves; as a bridge to other kinds of academic writing; and as a window into students’ developing language. Readers will find applications for ESL, bilingual education, composition, writing across the curriculum, writing research, literacy development, and deaf education. 1990, 154 pp., ISBN 0-939791-36-6, $11.95 ($9.95) plus postage and handling.

Ordering information is available from
Teachers of English to Speakers of Other Languages, Inc.
Publications Department
1600 Cameron Street, Suite 300
Alexandria, Virginia 22314-2705 USA
Telephone 703-836-0774
The TESOL Quarterly welcomes short evaluative reviews of print and nonprint publications relevant to TESOL professionals. Book notices may not exceed 500 words and must contain some discussion of the significance of the work in the context of current theory and practice in TESOL.


As recent TESOL Conference programs attest, the interest in literature in ESL classrooms, once greatly diminished, persists and even gains ground yearly. With this renewed interest in literature have come several anthologies of short literary pieces in English intended for use by ESL students. These anthologies usually include relevant materials and exercises to aid the ESL teacher in presenting and discussing the readings with nonnative speakers; lately, some volumes have been devoted specifically to demonstrating ways to use literature in ESL classes. Bursting with creative and imaginative suggestions, *Literature in the Language Classroom* is a stunning addition to this list.

After a few pages of general introductory reflections on using literature in language classes, the authors provide approximately 70 pages of possible activities for use before, during, and after reading a piece of fiction. This section supplies numerous sample worksheets demonstrating exactly how the suggestions can be implemented.

In the longest and most interesting section of the book, some of these suggestions are applied to specific texts. The authors take readers through dozens of activities to study a novel (*Lord of the Flies*), plays (*Romeo and Juliet* and Albee’s *The Sandbox*), several short stories, and a selection of poems. The end material includes two appendixes, one to help prepare students for questions on literature examinations and the other to provide teachers with a list of additional possible titles.

The authors agree with the current belief that pleasure reading provides the kind of language input that allows students to acquire enormous amounts of language without direct study of grammar, vocabulary, or discourse. For this reason the suggested activities are designed to promote the students' engagement with the most human aspects of a piece of literature rather than to provide linguistic practice or literary instruction. Thus, the authors neither belabor the linguistic benefits of reading fiction nor even mention exegetic approaches that rely on strictly literary apparatus such as the study of plot, characterization, or theme. Instead, students are encouraged to explore personal responses to a piece of literature: They are carefully guided toward seeing what the possible
range of responses might be, while being spared the frustrating problems
created when teachers find and impose interpretations that are completely
opaque to students. These activities build students’ confidence in their own
personal reactions to a piece of literature. At the same time they also
discourage, purely idiosyncratic interpretations by constantly calling for
students to use quotations and direct references to the text.

It is impossible to summarize the numerous fascinating and stimulating
suggestions offered in this book. Those of us who use literature in ESL
classrooms already know the pleasure it can provide both students and
teachers. Yet, even for novice teachers, Literature in the Language
Classroom is sure to inspire eager anticipation of the next opportunity to
teach literature and to experiment with Collie and Slater’s excellent
suggestions.

ILONA LEKI
University of Tennessee

Whole Language Strategies for ESL Students. Gail Heald-Taylor. San

Whole Language Strategies for ESL Students is the first book in a series
planned to explore issues concerning the nature and the development of
language and literacy. Edited by Jim Cummins, Sharon Lapkin, and
Merrill Swain, this series aims to provide educators as well as parents
information in a straightforward and accessible form within a theoretical
framework.

With this volume, the author and editors have succeeded in this goal.
This paperback handbook for primary grade (K-3) learners is a joy to read.
While the overall format facilitates ease of comprehension, the large clear
print with ample space between sections enhances ease of reading. A
chapter on the whole language philosophy opens the book, describing the
developmental language model and providing excellent references to
current research in language acquisition. This section concludes with a
discussion of specific benefits to ESL learners.

The next section offers an overview of whole language strategies
(highlighted and outlined in later sections) including dictation, literature,
writing process, themes, and evaluation. Each section is presented in two
columns. The left side describes the strategies for the general population
and is labeled “General.” The right side of each section describes the strat-
egies for the ESL student and is labeled “Implications for the ESL
Student.” Both columns are peppered with examples of each strategy
being presented. This format renders the handbook invaluable for the
mainstream teacher who has limited English proficient (LEP) students in
class, and for the ESL teacher or administrator who is responsible for the
professional development of mainstream teachers. For those ESL teachers
who note that they have been using many of these whole language strate-
gies for years, this handbook provides a clear theoretical framework.
One very practical section deals with the organization of whole language strategies and is filled with information helpful both to administrators and teachers. A daily timetable based on a language-arts time of 80 minutes or more per day is presented. Also provided is a weekly timetable for teachers who see students on a pullout basis for a half hour each day.

In addition to anecdotal records, dictation books, and writing samples, an “ESL Language Behavior Inventory” is presented in the evaluation section. This inventory, which can be duplicated for each student, provides a means of documenting language growth in listening, speaking, reading, and writing.

There is an excellent bibliography of trade books. These books, referred to throughout the volume, will not always be easily located by teachers in the U.S. because they are published in London or Toronto. A fine bibliography of professional references concludes this outstanding handbook.

The holistic direction in language pedagogy realized in, for example, the Natural Approach and the Whole Language Approach, has attracted a great deal of attention in recent years. This is the first easy-to-read, straightforward presentation of whole language theory and practice that is commercially available. This book is highly recommended for all teachers and administrators who work with second language learners.

FAY PALLEN
PS 26, District 2, Manhattan, NY


John Higgins, a leading figure in the field of computer-assisted language learning (CALL), has assembled a career’s worth of thought and reflection in his recent book, Language, Learners and Computers. In this short text, he puts forth his views regarding language learning while advocating his vision of how the computer might contribute to the acquisition of language. Having provided this background, Higgins furnishes a detailed rationale for what he feels is the proper perspective on the computer and its role in teaching and learning. Higgins guides us through his personal meditations on diverse language-related issues. His agenda throughout is to lay the foundation for his vision of the computer as “pedagogue” (p. 14) versus the more commonly shared perception of the computer as “magister” (p. 12). This compelling distinction has served as the basis of much of Higgins’ writing, materials development, and practice.

All discussion concerning the role of the computer in language learning, then, is knit to this magister-pedagogue analogy: the magister as one who possesses, and thus controls, the flow of knowledge versus the pedagogue as one who provides challenges and requisite information when asked to do so. Higgins contends that the computer can best serve language learners
in the more subservient role: The machine’s inherent domain is as tool, as slave, as something over which students can exert a certain amount of power and control. The computer is something to think with, not answer to. In contrast, the magisterial role, the role that the bulk of commercial language teaching software imposes, then, becomes a throwback to a style of instruction that, in the age of communicative language teaching, has a limited rationale.

The computer, according to Higgins, can best play the role of a thoroughly obedient slave who serves as a partner while one plays a game or solves a problem, something that a human is not likely to do well or felicitously. Higgins sees the role of the other player or partner as suiting the computer with its attributes of stupidity and blind obedience.

Students can use these very limitations of the machine for exploratory problem solving. Higgins uses the term exploratory to describe computer-based activities in which students actively debug language-generator programs that fail to accommodate a given set of rules and structures. In practice, students force the machine to make a mistake when generating sentences in the target language in order to test what the software can handle and is not able to handle.

The teacher’s role in such activities, says Higgins, is to set tasks and to “hover” (p. 26) as nonmagisterially as possible while students work in groups. This, says Higgins, can lead students to rely on reference books and to reflect on and solve a problem while calling on individual or collective knowledge bases. The teacher does act as a guide, seeking to “detrain” (p. 27) students who are likely to be accustomed to learning magisterially.

There persists throughout this text an intuitive appeal for the analogous distinction—master versus slave—which, if one manages to become indoctrinated in the first chapters, will allow one to overlook the author’s tendency to oversimplify a number of issues. Treatment of these language-learning issues becomes less convincing the less one is compelled by this distinction. The text’s chief mission, however, is to provide a new way of thinking about the computer in language teaching and learning, and it does present a cogent rationale for perceiving the computer as more than a task-master.

CARLA MESKILL
University of Massachusetts at Boston


The preface to this text states that the author’s intention is for Computers in the Language Classroom to serve as a “gentle introduction to the uses of computers in education” (p. i). In many respects, this presentation of material to teachers and teachers-in-training new to computers is gentle in
terms of ready accessibility and sensitivity to the noninitiated reader’s needs. But do not let the term gentle mislead you. This introduction to computers in language teaching is remarkably powerful.

Unlike texts written with a similar introductory purpose, here discussion of key issues in computer-assisted instruction (CAI) for language instruction is centered around ESL courseware designed, implemented, and field-tested by the author. Although the sophistication of this model software is now dated, the result of consistent reference to the how and why of a single language teaching product throughout each unit contributes to the text’s overall unity and clarity.

The author’s enthusiasm for the subject of computer-assisted instruction is engaging and at the same time nicely tempered by treatment of the realistic expectations and limitations of the technology in the field of language teaching. Hertz’s prevailing premise is that computers will (and are) radically changing our teaching lifestyle: from streamlining clerical work (recordkeeping and making handouts, for example) to taking over a great deal of tedious tutorial and testing work. Using the computer to simplify such work leaves the teacher with more time to be creative, to work on what teachers do best: develop exciting classroom activities and interact with students.

This text is dense with information. Useful lists of sources and reader worksheets are included. The main areas covered in the nine chapters include definitions of key concepts and terminology, discussion of applications, issues in the selection of software, hardware and peripherals, establishing a computer lab in your teaching institution, an introduction to programming and computer-assisted management as well as a stimulating chapter concerning the future of computers in education. These chapters are by no means as dry as the topics may suggest. They are, rather, practical, straightforward, easy to follow, and accompanied by intelligent discussion.

Every aspect of the text is consistent with its goal of initiating the language teacher into the field of computer-assisted instruction in language learning. The author does not oversimplify; he treats thoroughly and intelligibly issues often clouded by computer anxiety and lack of gentle training in the field. For the teacher who has been meaning to begin working with computers, this book can serve as a good guide. In all, *Computers in the Language Classroom* is an excellent, practical introductory text.

**CARLA MESKILL**

*University of Massachusetts at Boston*


This 16-chapter book is a pedagogical reference grammar intended for intermediate to advanced learners of English. Its purpose is to provide the...
reader with a description of English grammar that can also be used as a reference guide. The underlying assumption is stated in the rationale: Many nonnative speakers approach the study of English already equipped with a “grammatical consciousness” (p. vii), which is the result of their learning background.

The book is organized as a systematic (though not sequenced) linguistic description of the parts of speech (Chapters 2 through 9). It also includes explanations of individual grammatical features that usually present difficulties for students of English. Thus, it contains information at different levels, varying from relatively simple points to highly complex problems. Chapter 1, for example, deals with the sentence. After a few comments on word order and phrase structures, the author describes simple and complex sentence patterns, which provide a segue to discussion of indirect object movement, notions of dominance, relative clauses, adverbial clauses, and participial constructions. Description and problematic areas of the verb system are consistently and effectively treated in the last seven chapters.

A clear index and an extensive cross-referencing system help the reader find and retrieve all the information provided for each grammatical feature. Areas particularly problematic for the nonnative speaker are followed by a useful appendix; this includes lists of such things as phrasal verbs, prepositions, verbs requiring special constructions, and more detailed information about the topics discussed in the main body of the grammar. Explanatory and introductory material is in black, while examples are in blue.

Technical terminology is clearly defined and consists primarily of the traditional grammatical terms used in textbooks addressed to nonspecialists. Structures are first dealt with in terms of forms and then with regard to their use. Register variations and stylistic differences are mentioned where relevant. Although this book emphasizes usage in British English, it also provides some information about American English.

Linguistic description draws predominantly from traditional grammatical theory but also from updated research into grammar and discourse analysis. Because of the author’s avoidance of complex terminology and long descriptions, the reader is often provided with simplified versions of more detailed and sophisticated analyses originally devised by and for specialists. The resulting descriptions are accurate and appropriate to the level of the audience addressed.

The greatest strength of the book is its focus on description rather than prescription, on use rather than norm. Furthermore, Alexander presents the learner with a realistic description of a functionally differentiated, nonmonolithic language that can nonetheless be approached in a systematic way.

I would recommend the Longman English Grammar as a reference guide to be used both by the foreign and the second language learner of English at high-intermediate and advanced levels, on condition that both groups of learners have developed effective grammatical consciousness.
PAOLA SINIGAGLIA PUTTI  
*University of California, Los Angeles*


Although not concerned with style, rhetoric, or usage, *The Origin of Writing* does focus on a philosophical and historical perspective relevant to language teachers. For those of us whose language experience has been exclusively with the Indo-European family, it is hardly surprising to find that we have been directed toward an ethnocentric view of the origin of writing. Consider the relationship between ideas logically organized and set down on paper, and those techniques developed thousands of years ago (which we take for granted today) for the setting-down of those thoughts, whether on the level of alphabet, syllable, or ideogram: It would seem that if we are correct in perceiving that different types of languages have different ways to express ideas orally (English is not a tone language; Siswati has eight declension for its nouns), then perhaps something is askew in the concept that the alphabet supplies the basis for the highest form of written expression. In fact, the author suggests the possibility that the future of writing may bypass speech completely.

Harris lists some of the misconceptions of modern scholarship, chief among them that writing is an extension of speech or a representation of speech. He complains about an encyclopedia definition of writing as “the use of letters, symbols or other conventional characters, for the recording by visible means of significant sounds,” a view endorsed by major theorists from Saussure and Bloomfield to current thinkers. Harris cites as complementary oversimplifications the ideas that speech is the pronunciation of written forms, while writing is merely a way of setting down speech.

The author discusses the “evolutionary fallacy”: The idea of progression from embryo-writing through pictography to ideography, and then to phonetic writing of some kind, with the rebus believed to be a common transition device. As counterevidence, he points out that the rebus is simply arbitrary, with no possibility of systematization. Similarly, Harris is not impressed with theories of phonetic iconicity, which claim that written characters can be interpreted as articulatory diagrams relevant to the pronunciation in question, even though /o/ is the sound produced by unobstructed rounded lips, and is written as if those lips were being drawn. He also takes issue with the thesis that writing evolved from pictorial representation, and suggests that counting, and therefore graphic representation of numbers and types of objects counted, might have preceded the writing of speech.
The author denies the claim that writing exists solely and simply to set down what people say; the great significance about the invention of writing, to Harris, is the idea of “graphic communication as a mode of communication sui generis, adaptable to any particular communicational purpose desired.” He stresses that Saussure was “right to insist that speech was just one of many possible manifestations of a higher human faculty—that semiological faculty governing the creation of signs, which is ‘la faculté linguistique par excellence’” (p. 158).

Some of this is rather heavy going, but I found it generally worthwhile reading; I enjoyed the fascinating ruminations about American Indian personal names, the relevance of Platonic concepts of reality, and prehistoric Mesopotamian accounting systems. The real relevance to my life as a teacher comes with the realization that this view of the origin of writing connects directly with our understanding of writing as a separate and different activity from speaking. I think of my teaching of writing as facilitating a different kind of thinking and organizing; to find this view reflected, albeit on a different and far more sophisticated level, rewarded the effort.

ELAINE SELIPSKY

Intensive English Language Institute, Seattle Pacific University
Attitudes Toward “Cheating” Behaviors in the ESL Classroom

PHYLLIS KUEHN, DOUGLAS J. STANWYCK, and C. L. HOLLAND
Georgia State University

When ESL teachers encounter behaviors reflecting cultural value differences that they do not recognize, their most likely tendency is to judge those behaviors—and to characterize the students who engage in them—in terms of their own cultural value system. In an informal survey at Georgia State University, seven of nine university ESL instructors answered in the affirmative when asked whether international students cheat more frequently than U.S. students. Several years of observational data collected during test administration by one of the authors also provided evidence of attitude and behavioral differences between international students and U.S. teacher expectations. This perceived difference in behavior precipitates conflict in testing and classroom situations even for these rather culturally sensitive teachers.

In defining what constitutes cheating, students from different cultures apparently manifest differences of opinion and behavior in contrast to their U.S. student counterparts, not simply because of differences in role definition for student and teacher, but also because of more fundamental differences in embedded values of their respective cultures. In this regard Burton (1976) wrote, “Within the competitive Anglo-American culture, letting another copy one’s answers is cheating; but in Mexican and other cultures . . . [it] would simply be required altruism” (p. 196). Bagnole (1977), writing about the Arab kinship culture, stated that “cheating which does occur may contain additional cultural influences, not least of which is the compelling desire to help a friend” (p. 39).

The purpose of this study was to investigate the differences in attitudes of international students and U.S. students toward classroom cheating behaviors. A questionnaire was developed for the purpose of eliciting self-reports of attitudes toward three cheating behaviors: using crib notes, copying from another’s test, and allowing another person to copy, all in
hypothetical test situations. Given a scenario involving one of the listed behaviors, subjects were asked to indicate their judgment of the behavior by choosing one of six categories on each of two Likert scales anchored with right/wrong and moral/immoral. In addition, the act of refusing to allow someone to copy answers on a test was rated on scales of moral/immoral, loyal/disloyal, and friendly/unfriendly. In each situation students were asked to provide written comments that described the behavior in question.

The prototype questionnaire was developed in English, then translated into Spanish and Arabic, back-translated to check for accuracy, and retranslated after editing. Questionnaires were administered in classroom groups with institutional authority figures absent; all responses were anonymous. Responses were analyzed for 42 U.S. undergraduates, and 38 Spanish-speaking and 26 Arabic-speaking ESL students.

The right/wrong and moral/immoral response data were analyzed separately with two-factor (nationality-by-behavior) split-plot designs using both univariate and MANOVA approaches. The mean responses are plotted in Figures 1 and 2. Overall, all three groups rated the three behaviors on the wrong and immoral ends of the scale. On the right/wrong scale (Figure 1), there was not a significant group-by-behavior interaction (Wilks' $\lambda = .935$, approximate $F[4, 216] = 1.84$, $p = .122$), nor was there a

FIGURE 1
Right/Wrong Scale Mean Ratings of "Cheating" Behaviors
significant overall group difference in the ratings (approximate $F[2, 109] = 2.62, p = .078$). There was a significant effect for the "cheating" behaviors (Wilks' $\lambda = .912$, approximate $F[2, 108] = 5.24, p = .007$), indicating that they were not considered by the groups as equally wrong. In general, using crib notes was rated the most "wrong," and allowing another to copy from one's paper was rated as less wrong by the groups.

Ratings of the same behaviors on the moral/immoral scale (Figure 2) differed somewhat by group, although all three groups rated the behaviors at the immoral end of the scale. The nationality-by-behavior interaction was significant (Wilks' $\lambda = .908$, approximate $F[4, 204] = 2.52, p = .049$); the three groups perceive the behaviors as differentially immoral. Spanish speakers rated use of crib notes as most immoral while, in contrast, Arabic speakers rated allowing copying as most immoral. Overall, the group differences were significant (approximate $F[2, 103] = 5.76, p = .004$), but the "cheating" behaviors were not rated as significantly different on the moral/immoral scale overall (Wilks' $\lambda = .991$, approximate $F[2, 102] = .49, p = .615$). Although the groups all rated the behaviors toward the wrong and immoral ends of the scale, Spanish-speaker ratings were generally more neutral (toward the middle of the scale) than were those of the other two groups.

FIGURE 2
Moral/Immoral Scale: Mean Ratings of "Cheating" Behaviors

![Graph showing mean ratings of cheating behaviors across groups.]
Analyses of the written comments describing the behaviors provide some insights into the differences among the groups in their perceptions of the three behaviors. More than any other group, U.S. students categorized all three behaviors as cheating. Arabic and Spanish speakers tended to describe the behaviors as “dishonest” instead of as “cheating.” No Spanish speakers used the word cheating to describe allowing someone to copy.

Another class of comments referred to the ultimate harm done when students rely on these behaviors. Spanish (28%) and Arabic (27%) speakers (as compared to 5% of U.S. students) noted that the cheating behaviors go on, but indicated that what is wrong with them is the long-term problems associated with reliance on cheating, not that the behaviors themselves are inherently wrong or immoral. As one Spanish speaker wrote, “If he does it every day, very bad and harmful to his studies. If he does it on occasion because he couldn’t study, it’s understandable.”

In comments concerning allowing another to copy one’s work, Spanish speakers showed a difference in attitude compared to the other two groups; their comments (e.g., “I don’t care, I think the problem is for the person who copies”; and “Fernando loses nothing by letting Pedro copy from his exam”) are in contrast to those of 23% of the U.S. students and 18% of the Arabic speakers, who categorized allowing another to copy as cheating. The final comments by U.S. students revealed great concern about being in the position of having someone ask to copy their work. This is probably a reflection of the competitive nature of U.S. education as compared to the cooperative efforts described by Burton (1976) that are found in other cultures. Examples of the U.S. student comments about copying and allowing copying were “It was unfair for Fred to want to be graded on Steven’s knowledge and work,” “Fred took unfair advantage of his friend,” and “Fred wasn’t being fair—he put his friend on the spot.” This U.S. discomfort with the sharing of knowledge was also reflected in comments about the refusal of one student to allow another student to copy from his paper. U.S. students characterized this refusal as much more unfriendly than did the other two groups. The Arabic and Spanish speakers noted in comments that ultimately refusal is friendly since a good friend should be concerned about the long-term welfare of his friend. Their comments included “[Refusal is friendly] because this way the other student will try to study himself and this will make him sure of himself,” and “If Fernando allows him to copy, he will do him an injustice and this is not being a true friend.” Although the cooperative behavior is acknowledged, the ultimate harm in it is of more concern to the ESL students while the discomfort of such actions in the competitive context of U.S. education is of greater concern to the U.S. students.

The statistical results indicate that the Spanish speakers are less negative in their ratings of the “cheating” behaviors than are U.S. students or Arabic speakers; however, the nature of the differences is more complex than these data indicate. We can conclude from both the rating scale and comment data that there are attitudinal and behavioral differences between U.S. and international students, and that such differences cannot
legitimately be judged in terms of the values of this culture. What is cheating in one culture may have an entirely different value in another. As noted in the initial survey of ESL teachers, the interplay of these cultural differences can create problems in the ESL classroom. In their role as socializers of nonnative-speaking students into the academic culture, ESL teachers need to understand attitudinal differences in order to interpret the classroom behaviors they observe and to facilitate the transition of ESL students into the academic culture.

REFERENCES


Author’s Address: As of fall 1990, c/o Kuehn, School of Education and Human Development, California State University, Fresno, CA 93740.

Add ESL Students’ Management of Dialogue Journal Conversation

MARTHA R. DOLLY
Frostburg State University

Recent second language acquisition research suggests that conversational interaction, which involves two-way tasks that allow for negotiation of meaning, appears to be the key to developing the ability to participate successfully in oral interactions with native speakers (Long, 1983; Long & Porter, 1985); it is therefore important to explore L2 learners’ interfactional roles. The study described here extends previous investigation of another form of native-speaker (NS)/nonnative-speaker (NNS) interaction (e.g., Kreeft, Shuy, Staton, Reed, & Morroy, 1984) by analyzing the patterns of NNSs in conversation management of dialogue journal writing.

THE STUDY

The subjects were 12 adult ESL students (nine males and three females, representing seven first languages) enrolled in the intensive English program at Indiana University of Pennsylvania’s American Language Institute during the summer of 1985. All 12 volunteered to engage regularly in written conversation, known as interaction writing or dialogue journal writing, with the same NS teacher, as a supplement to their coursework. (In this report, though the terms NS and NNS are used, the focus is on the journal partners as student and teacher.)
Dialogue writing is similar to oral interaction in that, as the partners exchange the journal on a regular basis, each must participate to some degree in advancing and repairing the conversation. As in oral conversation, each partner may introduce, extend, and modify topics, as well as repair communication breakdowns. However, patterns of turn taking differ from those of oral conversation in that each journal partner responds to an entire entry (or turn) consisting of many moves, rather than to each move as it occurs. Entries grow out of each partner’s desire to share information and opinions unknown to the other, and success often requires negotiation of meaning. Dialogue writing thus qualifies as a two-way task, as described by Doughty and Pica (1986).

The conceptual framework and coding system devised for this study are based on work in the structure of oral conversation by Wells (1981), Long (1983), and others. These analytic systems were combined and adapted to describe the structure of dialogue journal discourse. Wells explains that discourse structure derives from each partner choosing, from a variety of options, an appropriate way of continuing. Choices made in the paradigmatic dimension (within each turn) relate syntagmatically to other turns because most moves carry some expectation of a follow-up move, thus providing a “structural link” (p. 28) to other turns. Wells’ system uses three types of moves: giving, soliciting, and acknowledging. A give carries less prospective force than a solicit, but more than an acknowledgment. In dialogue journal writing, each move allows for a reaction or response, but does not require it to the extent that oral conversation does; not all questions need be answered, nor all comments acknowledged.

As in oral conversation, each move in dialogue journal interaction serves to either advance or repair the conversation. A participant advances the conversation by initiating a new “topic framework” (Brown & Yule, 1983, p. 68), or by extending one previously introduced. Repairs interrupt conversation advancement to prevent or resolve a communication problem. Repairs in dialogue writing can be analyzed according to Long’s (1983) classification of repairs in oral conversation: clarification requests, confirmation requests, and comprehension checks. Repair moves, especially when followed up by the journal partner, involve negotiation of meaning and modify the interfactional structure of the written conversation in a manner similar to that of oral interaction.

Within this conceptual framework, a coding procedure for data analysis was developed. Each of the 260 student and teacher entries was analyzed in terms of its give and solicit moves. Each move was first coded as advancing or repairing. Advancing moves were then coded as initiating gives, reacting gives, responding gives, initiating solicits, and extending solicits. Repairs were coded according to Long’s (1983) classification system.

The purpose of the study was to investigate partners’ sharing of responsibility for conversation management. No two partners can assume equal responsibility for all conversational roles, but each must participate to some extent in introducing and extending topics and in handling communication problems. In this study, a conversation was said to be
reciprocal in a given move category when one partner used that move at least half as often as the other. Achieving reciprocity in a majority of the move categories suggests that a given pair of dialogue writers collaborates fully in maintaining a conversation.

RESULTS AND DISCUSSION

Five of the twelve conversations were reciprocal in a majority of the move categories. Nearly all journals were reciprocal in the initiating give category, but only four were reciprocal in initiating solicits, and only one in extending solicits. The teacher’s interfactional patterns remained fairly constant across all journals; the students tended to prefer some move types to others, though their conversational roles varied substantially.

Analysis of central tendency showed that the NNSs assumed most of the responsibility for initiating topics (57% of all student moves and 43% of teacher moves), and the NS, for extending topics; this is in contrast to Scarcella and Higa’s (1982) finding that, in oral conversation, NNSs prefer to comment on previously introduced topics familiar to them rather than initiate new ones. The NNS students’ fondness for topic initiation in dialogue writing may be due to the time lapse between turns, which reduces the awkwardness of moving from one topic to another. In addition, classroom dialogue writing encourages students (whether NS or NNS) to assume substantial responsibility for topic initiation rather than merely reacting or responding to teacher input (Kreeft et al., 1984); it would therefore be inappropriate for a teacher to initiate the majority of topics or to initiate topics without posing questions. Of the teacher’s initiating moves, about 30% were solicits, whereas of the students’ moves, nearly 90% were gives. The infrequency of student solicits does not necessarily suggest that NNSs hesitate to pose questions; as students, most of the NNSs took advantage of their opportunity to direct the conversation by providing information, relying on the teacher to encourage their topics by posing follow-up questions rather than formulating questions themselves in the course of introducing topics.

The greatest difference between student and teacher conversational roles was found in the extending solicit category (questions asked about previously initiated topics as opposed to questions asked in the course of initiating a topic). Extending solicits accounted for 14.8% of the teacher’s total moves, whereas the students devoted only 2.3% of their moves to extending a topic through questioning. Again, this difference probably reflects the teacher’s role as maintainer rather than initiator. In any case, when students did use extending solicits, it was usually to obtain needed information (e.g., “Where can we buy cheap and better quality things?”); the teacher, not needing such information, seemed to use extending solicits primarily as a tool for conversation management (e.g., “Did you buy anything at any of the stores you visited?”). Nonetheless, the NNS students might have benefited from extending the topics they initiated. An international student planning to continue his studies in Iowa begged his
partner, “Please tell me about tornado. I’m worrying about it? Please tell me!” However, he never acknowledged her reply or made any effort to pursue the topic.

Despite these trends and the NS’s consistent use of move types across journals, the degree and style of participation varied immensely among the students. One student seemed to be The Initiator because he introduced topics in rapid succession, posing many questions but rarely acknowledging the responses before rushing on. Another was dubbed The Solicitor because he posed more questions than the others. A student was labeled The Reactor because, especially during the first half of the journal, she preferred commenting on her partner’s material to introducing topics herself; later, however, she did introduce a few topics, including her “agony” over the marriage her mother was attempting to arrange for her. Another student, The Reporter, introduced many topics but rarely acknowledged the presence of his partner or sought reactions to his comments. Of all the students, the most flexible conversation partner was an older student, The Collaborator, who used all the move types his partner did (including extending solicits) in about the same proportions.

Perhaps the most unusual case was that of The Repairer, the student who frequently used comprehension checks (“I don’t know if you can understand me. I know that I am not very clear.”) to initiate negotiation of meaning.

A few other students initiated repair (such as the clarification request, “Please explain what you mean of ‘attitude’. ”) In most journals, however, repair was initiated only by the teacher. Repair appears to occur less frequently in interactive writing than in oral conversation, probably because understanding and being understood are less critical than in oral conversation (since not all comments and questions need to be addressed) and because each partner has time to ponder the other’s meaning. The NNSs who did initiate repair nearly always responded to the NS’s repair moves as well, whereas those who never initiated repair sometimes ignored their partner’s attempts to initiate repair. Gass and Varonis (1985) found that more negotiation work occurs in oral conversations between two NNSs than between a NS and a NNS. It may be that in the journal interactions, the NNSs were reluctant to admit a lack of comprehension and, because not all questions and comments had to be addressed, chose not to.

As analysis of conversational interaction replaces investigation of NS input, the NNS’s role in interaction becomes just as important as the NS’s. Dialogue journal writing provides a means of encouraging NNSs to experiment with conversational roles, freed from some of the pressures and constraints of oral conversation. The analytic system described here may allow teachers using dialogue journal writing with NNS students to gauge, formally or informally, the conversational styles of their students.
REFERENCES


Author’s Address: Department of English, Frostburg State University, Frostburg, MD 21532-1099
each in some of the world's most exciting and interesting countries in Asia, Africa, and Central Europe. Your expertise will be put to good use, and in return you'll receive invaluable experiences, memorable adventures, and unimaginable job satisfaction.

English Teacher Trainers are the professionals first requested by many of Peace Corps' new host countries. You'll be training the teachers of your host country, who will further the development and education process with the children and future leaders of their country.

See the world! Rather than travel as a tourist, why not live in a new country for two years? Speak a new language, experience a new culture, learn from and teach new neighbors, and discover a whole new way of life. Peace Corps makes it possible.

Opportunities like this define Peace Corps, and come with other benefits including free transportation, living expenses, medical and dental care, academic credit, student loan cancellation and/or deferment, and a $5400 re adjustment allowance.

Earn a new language, and a lot about yourself. You'll be challenged by places and people who will bring out new strengths and capabilities in you. You'll see why we call it "the toughest job you'll ever love" and why so many returned Volunteers call it the best experience of their lives.

Your experience, TESOL certification, or degree in Education or Linguistics can translate into numerous opportunities for you. If you are a U.S. citizen, and want to have the time of your life, contact your local Peace Corps recruiting office or write:

United States Peace Corps
Box 996
Washington, D.C. 20526 USA

(202) 606-3000
The Dirty Laundry of ESL Survey Research

JOY REID
University of Wyoming

For the novice researcher, the processes of design and implementation of a normed survey instrument for an ESL population hold many opportunities for error, and so for misleading reporting of results. This article traces the formulation and norming of one such questionnaire, constructed specifically for the self-reporting of ESL students on their preferred perceptual learning styles. This retrospective report provides description and comments on the problems of reliability and validity in survey design. It is presented here so that beginning researchers may gain some insight into the actual life of a research project, and to assist others working to design survey instruments for an ESL population.

Several years ago, I encountered almost by accident the CITE (Center for Innovative Teaching Experiences) Learning Styles Inventory (Babich, Burdine, Allbright, & Randol, 1975), a self-reporting instrument used by Kansas public schools to assist students—native speakers (NSs) of English—in identifying their preferred perceptual learning styles. It was a simply stated survey, normed on elementary and secondary NSs. My knowledge of learning styles was limited, but the idea of an instrument for such measuring, and for self-discovery, interested me. During the next several months, I investigated the available literature about learning styles; I found several additional survey instruments, and a wealth of information.

As I grew more knowledgeable, I decided to initiate a small pilot project, using the CITE instrument with approximately 120 nonnative-speaking (NNS) students in intensive ESL programs in Colorado. I enlisted the help of my colleagues in Colorado English language programs and, at the CoTESOL conference the following
autumn, reported on the results (Reid, 1982). As my interest in perceptual learning styles increased, I began recruiting volunteers—colleagues in English language programs across the country who agreed to administer a learning styles questionnaire to their ESL students (Reid, 1983). I then wrote a grant proposal for a larger project. As a consequence, I was awarded a year-long Graduate School Faculty Research Grant by Colorado State University (CSU) to (a) develop a valid and reliable questionnaire for ESL students, (b) administer that survey to about 1300 ESL students in approximately 45 intensive English language programs in the U. S., and (c) gather and interpret the resulting data. At that time, I considered the time line generous: I would put the survey together during the summer, mail it at the beginning of fall semester, retrieve and input the data by the end of that semester, and have the entire spring semester to do the necessary statistical analysis and interpretation of the data. What I did not realize at the time was how long and complex the process of norming a new survey instrument could be.

INITIAL SURVEY DESIGN

Even as I was completing the learning styles pilot project, I had begun to realize that a learning styles instrument such as the CITE survey, which had been normed on NSs (Babich & Randol, 1984), might not be as reliable or as valid with ESL students. This might be as a result of their different language and cultural backgrounds, or because of their attitudes toward and experiences (or lack of experiences) with self-reporting surveys. I therefore determined that I would develop a survey instrument for ESL students. Although my knowledge of survey design was extremely limited, I assumed that I would develop such a questionnaire through the following steps: I would

1. study available validated instruments;
2. select the statements from those survey instruments that I determined would fulfill my research design and expectations;
3. simplify the language as necessary for ESL students;
4. develop and organize the questionnaire; and
5. distribute it to the target audience

I was taking a statistics course at that time and presented my scheme for developing the survey instrument to the instructor, Professor Kenneth J. Berry. He emphasized the necessity of norming the survey by administering the pilot instrument to groups of students in order to make certain that the final instrument was
meeting my objectives. Establishing the reliability of my instru-
ment—either by estimating the correlation of scores obtained by the
same individual on different occasions (test-retest method) or with
different sets of equivalent items (split-half method)—would be
essential if I planned to achieve publishable results (K. Berry,
personal communication, 1984). Before I could do worthwhile
survey research, I needed to verify that the survey instrument

1. measured what I wanted it to measure (validity); and
2. measured each learning style equitably so that the research
could be successfully replicated (reliability)

Initially, I had decided that the survey would include six areas or
constructs: Visual, Auditory, Kinesthetic, Tactile, Group, and
Individual learning. (The last two are not actually perceptual
learning styles, but aspects of learning that I was also interested in
measuring.) From a review of the learning styles literature and of
existing learning style survey instruments, I decided that five
statements for each of the six constructs—a total of 30 statements—
would be sufficient. In a self-constructed instrument, asking
multiple questions in different forms that measure the same
construct (i.e., the same essential idea) helps to average out
response idiosyncrasies, thereby improving the validity of the
measurement process (Fowler, 1984). To assist me in the
development of the questionnaire, Berry provided me materials
concerning the successful formation of survey questions and the
process of randomly ordering the survey. He suggested that, in the
pilot testing of the survey, I should also double the number of
statements for each construct to a total of 60 statements. Using
twice as many statements as I intended in the final form of the
survey would afford me the opportunity to eliminate the least
successful statements and to select the “best five” in each category.
I was also to consult with second language experts to ensure that
those statements were clear and unbiased.

To examine the reliability of statements in the pilot survey, I was
referred to an educational statistician, Dr. Douglas Sjogren. During
that meeting we decided that, following the administration of the
survey instrument to a group of NSs and a group of ESL students,
we would first check the reliability of each pair of statements (see
Figure 1). That is, in order to see whether similar items were
eliciting the same responses from each respondent, we would look
to see whether, for example, a respondent who “strongly agreed”
with one statement in a pair would also “agree” or “strongly agree”
with the corresponding statement in the pair. A simple pair correla-
tion would show us whether the respondents were comprehending
the pairs, which, situated separately in the survey, should be eliciting similar responses such that the respondent was consistent in repeated measurements.

In addition Sjogren advised that, due to time limitations, I use the split-half method to assess the reliability of the survey instrument. In split-half reliability assessment, the test items are administered once to just one group of subjects; later, the items are divided into halves (ideally by random selection, but often by odd/even items). Each half is then treated as a subtest, and scores are obtained for each individual on comparable halves (Ary, Jacobs, & Razavich, 1979). For my survey, each construct of 10 statements would be viewed as a separate scale (or minitest); the 5 pairs of statements for each construct would be numbered, and then the odd “half-statements” (i.e., numbers 1, 3, 5, 7, and 9) would be compared with the even half-statements. The results of each subtest would then be correlated, using Pearson’s product-moment correlation, to obtain an estimate of reliability. The higher the correlation—which can range from −1 to +1—the more equivalent the halves, and the greater the reliability of the measure (Singleton, B. C. Straits, M. M. Straits, & McAllister, 1988). Sjogren (D. Sjogren, personal communication, 1984) judged that a split-half reliability coefficient of .70 would be indicative of a fairly reliable survey instrument.

As I later discovered, the major advantages of the split-half method are that it requires only one form of the test, thus avoiding any time lag in achieving results, and that the same physical and mental influences operate on the subjects. However, the split-half method systematically underestimates the reliability of the entire test, especially if each construct is measured by only a small number of items. In the case of my survey, for example, each construct comprised a set, with only 10 items (5 pairs) in each set. If one of my construct split-half reliability coefficients was .65, the actual reliability coefficient might be as high as .79 (Ary et al., 1979).

While I felt that I was learning a good deal about survey design and implementation, I was disappointed by the slowing of my timeline, when I agreed (a) to show the proposed survey to colleagues in our intensive language program and in the MATESL program and (b) to double the size of the survey. In addition, I decided to administer the survey to groups of ESL students in the CSU Intensive English Program and to a group of NSs at CSU, in order to compare results. I returned to the several original learning style instruments for NSs that I had encountered in my reading. Because those surveys had been normed previously, I found that, for the most part, the statements and questions were well designed. The vocabulary and sentence structures were clear and simple for the
NS population; each statement contained only one point for response; there were no “loaded guns,” no leading questions; there were no presumptive or potentially embarrassing questions; and the 5-point Likert scale provided sufficient response options: Strongly Agree, Agree, Undecided, Disagree, and Strongly Disagree (Fowler, 1984; Leedy, 1985). In some cases, the available statements seemed completely appropriate for my survey, while others required rewording, specifically for clarity and simplicity of language, to meet the needs of the ESL population.

After I had formulated a draft of the survey, I consulted with colleagues: teachers in the Intensive English Program and in the English Department’s MATESL program. These experts suggested some small (generally cosmetic) changes; based on their input, I made the necessary revisions. Then, using the procedures recommended in the survey research literature, I randomly selected the statements and put together the questionnaire. The resulting survey contained six constructs (Visual, Auditory, Kinesthetic, Tactile, Group, and Individual). Each construct consisted of five pairs of statements, with each pair stating a similar point in a slightly different way. The survey therefore consisted of 10 statements for each construct, a total of 60 statements. Figure 1 provides examples

FIGURE 1
Sample Paired Statements from the Original Learning Styles Survey

<table>
<thead>
<tr>
<th>Visual learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. When the teacher writes things on the chalkboard, I understand them better.</td>
</tr>
<tr>
<td>2. I learn better when the teacher writes things on the board.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Auditory learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I remember things I have heard better than things I have read.</td>
</tr>
<tr>
<td>2. I understand things I have heard better than things I have read.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Kinesthetic learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I prefer to learn by experience rather than by listening to lectures.</td>
</tr>
<tr>
<td>2. I learn better by experience than I do by listening to lectures.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tactile learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I learn more when I make something for a class project.</td>
</tr>
<tr>
<td>2. I enjoy making something for a class project.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I enjoy working on an assignment with two or three classmates.</td>
</tr>
<tr>
<td>2. When I am working on an assignment, I like working with two or three classmates.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Individual learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I learn best when I study alone.</td>
</tr>
<tr>
<td>2. When I study alone, I remember things better.</td>
</tr>
</tbody>
</table>
of statement pairs for each construct, although these did not appear side by side in the survey questionnaire.

I administered the questionnaire to more than 50 ESL students and to more than 50 NSs; the students included both graduate and undergraduate students at CSU who were in a variety of fields of study. Douglas Sjogren then helped me run a correlational study on the 60 responses in each survey from over 100 students. Using Pearson’s product-moment correlation, we checked the intercorrelation between each pair of statements in each construct; then, using the split-half method, we correlated each statement with the other statements in the construct. We examined the results (in the form of correlation coefficients) for the two groups: ESL students and native speakers. The results of the pair correlation coefficients are shown in Table 1.

Next, using the split-half method again, Sjogren and I examined the correlation coefficients for the interrelatedness of the 10 statements in each of the six constructs. Table 2 gives those results.

DISCUSSION OF INITIAL SURVEY RESULTS

As we examined the results of the analysis of my survey, some of the correlational results immediately demonstrated that some pairs of statements were directly not synonymous for either the NS or the ESL students:

Auditory

1. When I need to remember something, I say it aloud several times.
2. In class lectures, I learn better by listening to the teacher without taking notes.
(Pair correlation coefficients: NS = –.0298, ESL = .1320)

Other pairs of statements were startling in their discrepant NS and ESL correlation coefficients:

Visual

1. I remember instructions I have read better than instructions I have heard.
2. I understand better when I read the instructions instead of listening to someone tell me the instructions.
(Pair correlation coefficients: NS = .8031, ESL = .3832)
<table>
<thead>
<tr>
<th>Native speakers of English</th>
<th>ESL students</th>
</tr>
</thead>
</table>

**Visual learning**
1. .4238  
2. .3334*  
3. .5747  
4. .6031  
5. .5425*  
1. .3527  
2. .3037  
3. .8289  
4. .8532  
5. .2245

**Auditory learning**
1. .3629*  
2. .5422  
3. -.0298*  
4. .4343  
5. .8663  
1. .2524  
2. .2616  
3. .1320  
4. .1865  
5. .5967

**Kinesthetic learning**
1. .5725*  
2. .7303  
3. .7637*  
4. .7526  
5. .5987  
1. .5429  
2. .7644  
3. .0900  
4. .5155  
5. .3932

**Tactile learning**
1. .6795  
2. .5794*  
3. .6498*  
4. .8056  
5. .7394  
1. .4537  
2. .3927  
3. .1145  
4. .4403  
5. .5445

**Group learning**
1. .7322*  
2. .7252*  
3. .7443  
4. .5132  
5. .7826  
1. .5352**  
2. .4256**  
3. .6264  
4. .5445  
5. .7212

**Individual learning**
1. .7586  
2. .7039  
3. .7311  
4. .6278*  
5. .6364*  
1. .5415  
2. .6095  
3. .4537**  
4. .7244  
5. .3744**

*Note.* An asterisk indicates pairs that would be eliminated from a final questionnaire were this survey to be used with NSs. The pair with the lowest correlation in each construct, as well as any others whose wording did not seem appropriate to the construct, was eliminated. A double asterisk indicates pairs that were eliminated from the Group and Individual constructs in the final ESL survey because of their low correlation coefficients.
Correlation Coefficients for the Six Constructs in the Original Learning Styles Survey

<table>
<thead>
<tr>
<th>Construct</th>
<th>NS</th>
<th>ESL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual</td>
<td>.3988</td>
<td>.2158</td>
</tr>
<tr>
<td>Auditory</td>
<td>.4008</td>
<td>.3948</td>
</tr>
<tr>
<td>Kinesthetic</td>
<td>.5470</td>
<td>.3969</td>
</tr>
<tr>
<td>Tactile</td>
<td>.6235</td>
<td>.3475</td>
</tr>
<tr>
<td>Group</td>
<td>.6492</td>
<td>.5467</td>
</tr>
<tr>
<td>Individual</td>
<td>.6538</td>
<td>.6076</td>
</tr>
</tbody>
</table>

**Tactile**

1. I really like to build things.
2. When I build something, I remember what I have learned better.

(Pair correlation coefficients: NS = .8056, ESL = .4403)

We were not surprised to discover that the correlation coefficients for the NSs and the ESL students differed; after all, the NSs were more language proficient, generally more experienced in completing surveys, and culturally more homogeneous than the ESL students. What seemed more surprising was the range of differences; the NSs, quite simply, responded to the survey statements in a quantitatively (and so, perhaps, a qualitatively) different way. They found the questionnaire exceedingly easy to understand, and they generally responded consistently to pairs of statements and to the 10 statements within each construct. For example, if the respondent preferred visual learning, her responses to the 10 visual learning statements were usually Strongly Agree or Agree on the Likert scale. In fact, at the end of one of the surveys, a native speaker had written: “This this this this this questionnaire questionnaire questionnaire questionnaire is is is is repetitive repetitive repetitive repetitive repetitive.” Sjogren explained that, if I wanted to use the survey for NSs, I needed only remove those items that correlated least with the total construct score (see single-asterisk items in Table 1), in order for the construct correlation coefficient to rise substantially to about the .70 level. I would thus have a survey that was reliable for a NS population. Unfortunately, since my target population was NNSs, I abandoned these data.

In contrast to the consistency of the NS results, the correlation coefficients for the ESL survey responses were, generally,
substantially below those of the NSs. In fact, only the Group and Individual constructs would remain if the statements that correlated least well were eliminated. Double asterisks in Table 1 indicate the pairs eliminated from the Group and Individual constructs; also, one additional statement was eliminated from each of the Group and Individual constructs, leaving a total of five statements in each. In the other constructs, the ESL student responses showed that their perceptions of the statements within a construct differed; that is, responses were inconsistent, with some of the statements within a set marked Strongly Agree and others Disagree or, just as worrisome, Undecided. Most probably, the ESL respondents did not understand the language in many of the items; some of the statements were, at best, ambiguous, at worst incomprehensible. The results of such inconsistencies were clearly demonstrated in the extremely low coefficients for pair correlations and construct correlations.

As I reconstructed my norming procedures for this article, I became increasingly aware of the lack of information concerning absolute cut-off correlation coefficients for reliability in surveys. Despite the fact that .70 had originally been recommended as a minimum goal, I learned that assigning a single baseline correlation was not easy. Indeed, in a recent conversation, three statisticians confirmed that research methodologists differ in their opinions concerning the level of equivalence for a reliable survey (K. Berry, personal communication, 1989; M. Miller, personal communication, 1989; D. Sjogren, personal communication, 1989); generally speaking, a good reliability is one that is as good or better than the reliability of competing measures. Since no other normed learning styles instrument for ESL students had been published, I had no competing measures with which to compare my survey. In addition, both the length and the function of the test influence the level of acceptable reliability. The longer the test, the greater its reliability—unfortunately, each of my constructs had to be correlated as a separate test, and each of those constructs was very short. The function of my survey might also lower its reliability. Whereas tests of academic achievement (e.g., spelling) that can be measured objectively often have quite high reliability (.80 or .90), tests of personality variables, such as a self-reporting survey of perceptual learning styles, often have only moderate reliabilities (approximately .60; Ary et al., 1979). Finally, the reliability of an instrument is partly a function of the ability of the individuals who take the test; because of the diversity of my survey population—ESL students from many different countries and cultures, and with limited and varying experiences with survey instruments—the correlation coefficients that were slightly lower might still result in a reliable survey. In addition, Sjogren said that generally correlation
coefficients for constructs should also be at least .70 for adequate reliability of a survey instrument; however, the diversity of the test population as well as the use of the split-half method to measure reliability, which results in underestimated correlation coefficients, might make such coefficients difficult to achieve.

REVISING THE SURVEY

However, even if .60 was to be my baseline correlation coefficient, I had not achieved even that level with the ESL surveys. Sjogren articulated my worst fears: It would be necessary to redesign pairs of statements for the first four constructs of the survey and pilot test the new questionnaire with ESL students. Suddenly my time line became inadequate. Most important, I did not yet have information to indicate what was wrong with the survey. I decided to look at the pairs of statements that had high correlation coefficients, to analyze what made those statements accessible to ESL students. Then I would confer with the true experts, ESL students, who could perhaps explain the failure of some of the statements to elicit consistent responses.

A dozen ESL students from different language and cultural backgrounds agreed to help, some from the intermediate and advanced levels of the Intensive English Program (my target population), some postadmission students from different graduate and undergraduate major fields, and still others who were studying in the MATESL program. These students generously spent many hours with me, going through every questionable statement on the survey, rewording and restructuring to render each statement clear and unambiguous. Many of the statements needed simpler verb tenses and less idiomatic vocabulary; others needed to have structural complexity reduced by removing the comparative dependent clause at the end of the statement. In other cases, a single difficult word might preclude clear communication; sometimes the cognitive load of a statement was too great for the respondents’ immediate understanding. Finally, we discovered that the smaller the number of sentence and vocabulary patterns we used, the better the chance of a comprehensible instrument; even though the statements then sounded repetitious, they were clearer to the students. Some of the most successful pair statements contained the following:

1. I learn better when I . . . .
   I think better when I . . . .
2. When I [verb phrase] . . . I remember better.
   When I [verb phrase] . . . I understand better.
3. I like learning and/or working . . . .
   I enjoy learning and/or working . . . .
4. I prefer to learn . . . .
   I like to learn . . . .

Figure 2 provides examples of the kinds of changes we made, with the rationale for each change.

FIGURE 2
Revised Pair Statements for the learning Styles Survey

Visual (Change statement of multiple focus to single focus.)
Original: When I learn something new, I like to learn about it by seeing it in a book, a photo, or a diagram.
Revised: When I learn something new, I like to learn about it by seeing it in a photo.

Auditory (Remove idiomatic what to do and the last comparative phrase.)
Original: I understand instructions better when the teacher tells me what to do than if I must read the instructions.
Revised: When the teacher tells me the instructions, I understand better.

Kinesthetic (Remove difficult vocabulary words add explicit direct object in the main clause.)
Original: I learn best when I can participate in related activities.
Revised: I understand things better in class when I can participate in activities.

Tactile (Substitute understand for like; add a comparative word.)
Original: I really like to build things.
Revised: I understand things better when I build them.

Following an intensive 2-week revision of the instrument, I administered the survey to 50 ESL students. This survey was shorter: After considerable discussion with Sjogren, I did not retest the already reliable Group and Individual constructs, and I administered only four pairs of Kinesthetic statements and one Tactile pair of statements to the new population. Following the administration of the questionnaire, Sjogren and I again went through the pair correlations and the split-half correlational analysis. Except for the Visual construct, the pair correlation coefficients were, indeed, substantially higher, although the construct correlation coefficients remained low. Table 3 shows the pair and the construct results for the revised questionnaire.

With the exception of the Visual construct, the results of the correlational study were successful. We could remove the items/statements in each construct that correlated least well with the total
score; the construct correlation coefficient for the remaining five statements would then fall within the acceptable range for a reliable survey.

As for the Visual construct, the construct correlation coefficient was even lower than before! My time line was overextended and fall classes were about to begin. In spite of this, Sjogren was excited about the discrepant correlations because, as he looked at the statements involved, he discovered that a basic ambiguity seemed to exist, for my population, between statements that concerned visual text reading and visual nontext reading. In other words, while most of the statements focused on reading text, several involved the “reading” of charts, graphs, and photos: ESL students evidently perceived reading language as being quite different from reading illustrations. His suggestion was that I focus on this interesting point: Redesign just the Visual construct, dividing it into two separate sets—visual-text and visual-illustration—then readminister the questionnaire to 50 more ESL students.

Sjogren also suggested that I pretest the survey once again, this time broadly, with several hundred students, and do an item analysis to see which individual items contributed the most to each construct. In retrospect, I should have followed that advice; the potential for a more reliable survey instrument, additional interesting data, and essential information about ESL students responding to surveys was great. However, I felt that I did not have the time or the resources to do yet another split-half correlational study, much less a broad-based item analysis pretest; I indicated that I wanted to use the survey as it stood if we could “fix” the Visual construct. After the illustration (nontext) statements were

---

<table>
<thead>
<tr>
<th>Pair correlations</th>
<th>Visual</th>
<th>Auditory</th>
<th>Kinesthetic</th>
<th>Tactile</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. .6120</td>
<td>1. .4436*</td>
<td>1. .3857*</td>
<td>1. .5779</td>
<td></td>
</tr>
<tr>
<td>2. .2396*</td>
<td>2. .5702</td>
<td>2. .7750</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. .3728*</td>
<td>3. .8471</td>
<td>3. .3296*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. .8005</td>
<td>4. .1174*</td>
<td>4. .4114</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. .7039</td>
<td>5. .5302</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Construct correlations</th>
<th>Visual</th>
<th>Auditory</th>
<th>Kinesthetic</th>
<th>Tactile</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0801</td>
<td>0.3155</td>
<td>0.4273</td>
<td>0.4698</td>
<td></td>
</tr>
</tbody>
</table>

*An asterisk indicates pairs with low correlations that were eliminated from the final survey instrument.*
removed, the construct correlation coefficient fell within the acceptable range for a reliable instrument.

By the end of that week, I had the survey statement randomized, appropriately laid out, copied, and ready to be mailed to the 43 English language programs that had agreed to participate in the study. The results of the survey were published in the *TESOL Quarterly* (Reid, 1987).

**IMPLICATIONS**

Why air this dirty laundry? First, because I believe that I was as committed as one could be to carrying out careful, responsible research, and as typically naive as are many novice researchers, about the problems I would encounter. Second, in the years since my first empirical endeavor, I have realized how informative such research can be for the researcher; indeed, research is a discovery process (Cooper, 1982). It must be seen as a context in which to develop theoretical knowledge (McGuire, 1983). Finally, I discovered that merely reading published research does not allow the new researcher insight into the extent to which most research (survey, ethnographic, empirical, or historical) is cleaned up and manicured for publication. A new researcher might well think, as did I, that simply deciding on a research design will result in clear, clean implementation of the methodology and in unambiguous results.

Of course, I do not present my study as an ideal. The construct coefficients, for instance, could have been higher, and my survey more reliable, had I followed the counsel of my advisers and worked a little longer on the item analysis. Furthermore, the information might have been more compelling had I divided the Visual learning construct into two parts. Instead, I have deliberately washed my own research laundry in public because I have decided, in the intervening years since my research was published, that my experiences in developing a survey instrument for ESL students might well be of interest to others. Nor do I present the norming of my survey as an example of my professional acumen; indeed, without two mentors, whom I encountered serendipitously and consulted casually, my research would have been inconclusive and misleading at best. I can say, however, that since that investigation, I have taken several statistics courses and have become interested in the processes as well as the products of research. I now recognize the need for expert help; as a result, I never embark on a research project without the detailed advice of a statistician.

The dirty laundry of this particular article must also be washed: My greatest oversight in carrying out the learning styles research
was that I almost immediately threw out all of the raw data used to
norm my survey instrument. As I sought, five years later, to
reconstruct the process of norming the survey, I began digging
through files, rescuing scraps of paper. I met, once again, with
Berry and Sjogren (the latter now retired) in order to use our
combined memories to assure accuracy for this article. Alas, I
discovered that I no longer had the numerical construct correlation
coefficients for the final draft of the questionnaire, although they
did once exist and were sufficiently reliable. Unfortunately, I could
not reconstruct them, because the original surveys were gone. By
that time, I had invested much energy in the planning and drafting
of this article, and we agreed that the information was worth
publishing, even if the data were incomplete. My apologies to the
reader for not having this essential data set.

Perhaps a more important reason for washing the dirty laundry of
my learning styles research was that, in the years since I designed
and completed my study, I have become aware of the increasing
amount of research in TESOL that is statistically analyzed and
reported, some of it with baldly unnormed instruments or with
questionnaires that have been normed for NSs but not for ESL
students. I am wary of results based on unnormed instruments, and
of global interpretations of such data. Not only have I seen, first-
hand, an ESL population respond discrepantly to what I thought
was a simple, straight-forward survey, but I have also discovered,
through the data from my questionnaire, that ESL students from
different language and cultural backgrounds may also differ in the
ways they respond to normed surveys. For example; in the learning
style data I collected, NSs used the entire range of the 5-point Likert
scale in a reasonably consistent manner, while the Japanese students
tended to respond more toward the mean: That is, they responded
to the Strongly Agree and the Strongly Disagree categories only
rarely. This tendency prevented the Japanese language group from
displaying clearly defined, major learning style preferences, as did
subjects from some of the other language backgrounds represented
in the survey.

Finally, I believe that most research has some “dirty laundry” that
has been eliminated and/or concealed before publication. Readers
of research should, therefore, be aware of the possible pitfalls of re-
search design and implementation. More specifically, both re-
searchers and readers of research should be aware of the rather
significant problems involved in norming a survey instrument for an
ESL population, as well as the specific problems involved in using
a nonnormed instrument or a questionnaire that has been normed
only for NSs. As other researchers have used my survey (Dirksen,
1988), or translated it for their use (Melton, 1988), or adapted it for use in their investigations (Eliaison, 1989), I have become increasingly aware of the necessity of appropriately normed instruments in ESL research. Without normed instruments, replication will not be possible.

Of course, questionnaires that have not been normed for the population that is being studied may have a place—for example, in the ESL classroom. A nonnormed survey that helps students discover their stated preferences for study environment may be an excellent teaching tool. Student-developed surveys in oral communication classes can spark discussion and increase student participation. And the use of a survey such as the Meyers-Briggs Type Indicator for NSs might be useful as a teaching tool in raising student awareness and/or in cross-cultural training. In short, if the results of nonnormed surveys are used as consciousness-raising tools for ESL students, with the essential caveat that these instruments are not necessarily either reliable or valid, they may prove useful in the classroom.

However, the results of such surveys should never be used as the basis for curricular change; they should not be used for policy making or for institutional change. Indeed, to take such results as valid and reliable, and to use them accordingly, will prove a disservice to students, teachers, and researchers.

ACKNOWLEDGMENTS

My thanks to my original mentors, Kenneth J. Berry (Departments of Sociology and Statistics, Colorado State University) and Douglas Sjogren (Department of Education, Colorado State University), for their advice and assistance. In addition, I am grateful to Marinell Miller (Missouri LINC, University of Missouri) for her suggestions, and to Barbara Kroll (California State University, Northridge) for her generous editorial assistance.

REFERENCES


INFORMATION FOR CONTRIBUTORS

EDITORIAL POLICY

The TESOL Quarterly, a professional, refereed journal, encourages submission of previously unpublished articles on topics of significance to individuals concerned with the teaching of English as a second or foreign language and of standard English as a second dialect. As a publication which represents a variety of cross-disciplinary interests, both theoretical and practical, the Quarterly invites manuscripts on a wide range of topics, especially in the following areas:

1. psychology and sociology of language learning and teaching; issues in research and research methodology
2. curriculum design and development; instructional methods, materials, and techniques
3. testing and evaluation
4. professional preparation
5. language planning
6. professional standards

Because the Quarterly is committed to publishing manuscripts that contribute to bridging theory and practice in our profession, it particularly welcomes submissions drawing on relevant research (e.g., in anthropology, applied and theoretical linguistics, communication, education, English education [including reading and writing theory], psycholinguistics, psychology, first and second language acquisition, sociolinguistics, and sociology) and that address implications and applications of this research to issues in our profession. The Quarterly prefers that all submissions be written so that their content is accessible to a broad readership, including those individuals who may not have familiarity with the subject matter addressed.

GENERAL INFORMATION FOR AUTHORS

1. The TESOL Quarterly invites submissions in five categories:

   Full-length articles. Manuscripts should generally be no longer than 20 double-spaced pages. Submit three copies plus three copies of an informative abstract of not more than 200 words. To facilitate the blind review process, authors' names should appear only on a cover sheet, not on the title page; do not use running heads. Manuscripts should be submitted to the Editor of the TESOL Quarterly:

   Sandra Silberstein
   Department of English, GN-30
   University of Washington
   Seattle, WA 98195 U.S.A.

   Reviews. The TESOL Quarterly invites reviews of textbooks, scholarly works related to the profession, tests, other instructional materials (such as computer software, videotaped materials, and other nonprint materials),
and other journals concerned with issues relevant to our profession. Comparative reviews, which include a discussion of more than one publication, and review articles, which discuss materials in greater depth than in a typical review, are particularly welcome. Reviews should generally be no longer than five double-spaced pages, although comparative reviews or review articles may be somewhat longer. Until further notice, submit two copies of reviews to the Review Editor of the 

TESOL Quarterly:

Heidi Riggenbach
Department of English, GN-30
University of Washington
Seattle, WA 98195 U.S.A.

Book Notices. The TESOL Quarterly also welcomes short evaluative reviews. Book notices should provide a descriptive and evaluative summary of a recent publication (see preceding section for appropriate types of publications) and a brief discussion of the significance of the work in the context of current theory and practice in the relevant area(s) of TESOL. Submissions should range between 350 and 500 words; any submission that exceeds 500 words will be returned. Submit two copies of book notices to Heidi Riggenbach, Review Editor, at the address given above.

Brief Reports and Summaries. The TESOL Quarterly also invites short descriptions of completed work or work in progress on any aspect of theory and practice in our profession. Reports of work in the areas of curriculum and materials development, methodology, teaching, testing, teacher preparation, and administration are encouraged, as are reports of research projects of a pilot nature or that focus on topics of specialized interest. In all cases, the discussion of issues should be supported by empirical evidence, collected through either formal or informal investigation, and should be grounded with references to similar or related work. Manuscripts should summarize key concepts and results in a manner that will make the research accessible to our diverse readership. Although all reports and summaries submitted to this section will be considered, preference will be given to manuscripts of three to seven double-spaced pages (including references and notes). Longer articles do not appear in this section and should be submitted to the Editor of the TESOL Quarterly for review. Send two copies of reports and summaries to the Editor of the Brief Reports and Summaries section:

Gail Weinstein-Shr
Reading and Writing Program
School of Education
Furcolo Hall
University of Massachusetts
Amherst, MA 01003-0096

The Forum. The TESOL Quarterly welcomes comments and reactions from readers regarding specific aspects or practices of our profession.
Responses to published articles and reviews are also welcome. Contributions to The Forum should generally be no longer than five double-spaced pages. Submit two copies to the Editor of the TESOL Quarterly at the address given above.

2. All submissions to the Quarterly should conform to the requirements of the Publication Manual of the American Psychological Association (Third Edition), which can be obtained from the Order Department, American Psychological Association, 1200 Seventeenth Street, NW, Washington, DC 20036. The Publication Manual is also available in many libraries and bookstores.

3. All submissions to the TESOL Quarterly should be accompanied by a cover letter which includes a full mailing address and both a daytime and an evening telephone number. Where available, include an electronic mail address and fax number.

4. Authors of full-length articles should include two copies of a very brief biographical statement (in sentence form, maximum 50 words), plus any special notations or acknowledgments that they would like to have included. Double spacing should be used throughout.

5. The TESOL Quarterly provides 25 free reprints of published full-length articles and 10 reprints of material published in the Reviews, Brief Reports and Summaries, and The Forum sections.

6. Manuscripts submitted to the TESOL Quarterly cannot be returned to authors. Authors should be sure to keep a copy for themselves.

7. It is understood that manuscripts submitted to the TESOL Quarterly have not been previously published and are not under consideration for publication elsewhere.

8. It is the responsibility of the author(s) of a manuscript submitted to the TESOL Quarterly to indicate to the Editor the existence of any work already published (or under consideration for publication elsewhere) by the author(s) that is similar in content to that of the manuscript.

9. The Editor of the TESOL Quarterly reserves the right to make editorial changes in any manuscript accepted for publication to enhance clarity or style. The author will be consulted only if the editing has been substantial.

10. The views expressed by contributors to the TESOL Quarterly do not necessarily reflect those of the Editor, the Editorial Advisory Board, or TESOL. Material published in the Quarterly should not be construed to have the endorsement of TESOL.
Want to stimulate your students' reading? This book will help. Pass it on.

A WORLD OF BOOKS
An Annotated Reading List
For ESL/EFL Students
Second Edition, Revised

Dorothy S. Brown

Pass A World of Books on to intermediate and advanced students, and let them choose 20th century titles that appeal to their tastes.

Some entries are marked Y (for younger readers) and some are marked M (for mature readers). Appendices classify the fiction and nonfiction books according to setting, subject matter, and level of difficulty.

A World of Books will "help students who are fairly proficient in English, but not completely at home in the cultures of English-speaking countries, to find books that they can read with a maximum of pleasure and a minimum of frustration." (Dorothy S. Brown)

$6.00 members, $7.50 nonmembers. Plus 15% postage. All orders must be prepaid. VISA, MasterCard welcome. 6" x 9" softbound, 70 pp. ISBN 0-939791-32-3

Teachers of English to Speakers of Other Languages
1600 Cameron Street, Suite 300
Alexandria, Virginia 22314 USA
703-836-0774
PUBLICATIONS RECEIVED

Publishers are invited to send copies of their new materials to the TESOL Quarterly Review Editor, Heidi Riggenbach, University of Washington, at the address listed in the Information for Contributors section. Packages should be labeled REVIEW COPIES.

TESOL Quarterly readers are invited to contribute review articles and evaluative or comparative reviews for consideration for publication in the Review or Book Notice sections of the Quarterly. These should be sent to the TESOL Quarterly Review Editor, Heidi Riggenbach, University of Washington, at the address listed in the Information for Contributors section.

TESOL gratefully acknowledges receipt of the following publications.


PUBLICATIONS AVAILABLE
FROM THE TESOL CENTRAL OFFICE

Publications on a wide range of topics relevant to TESOL professionals are available from the TESOL Central Office. Many of these publications are offered at reduced rates to TESOL members. An annotated list of publications can be obtained by contacting TESOL Publications Department, Suite 300, 1600 Cameron Street, Alexandria, Virginia 22314 U.S.A. A TESOL Order Form is printed in every issue of the TESOL Quarterly and can be used to order any of these materials. Please note that prices inside parentheses are special discount prices for TESOL members only. Prices are subject to change without notice.

REFERENCE GUIDES

*Directory of Professional Preparation Programs in TESOL in the United States, 1989-1991.* Describes 245 college and university programs in TESOL leading to doctorate, master’s, or bachelor’s degrees, and certificates. Overview of state certification requirements. 1989. $20.00 ($15.00). ISBN 0-939791-35-8

*1987 TESOL Membership Directory.* A comprehensive resource of who’s who in TESOL. Both alphabetical and geographical listings of more than 11,000 institutional, commercial, and individual members. $5.00.


*Annotated Bibliography of ESL Materials.* Christine Aronis. Gathered from 76 publishers, the 636 ESL titles listed were designed to teach American English. Provides ESL teachers with an accurate and comprehensive list of titles available. 1983. Sale $5.00.


THEMATICS


*Students and Teachers Writing Together: Perspectives on Journal Writing.* Joy Kreeft Peyton, editor. Examines journal writing as a powerful stimulus for curriculum and teacher change; as a forum for discussing the meaning and form of literary texts and the journal texts themselves; and as a bridge to other kinds of academic writing. Applications for ESL, bilingual education, composition, writing across the curriculum, writing research, literacy development, and deaf education. 1990. $11.95 ($9.95). ISBN 0-939791-36-6
Research in Reading in English as a Second Language. Joanne Devine, Patricia L. Carrell, and David E. Eskey, editors. Views of reading as an interactive process involving reader and text variables, of the interplay of reading skills and general language proficiency, and of the differences in first and second language reading. 1987. $16.50 ($15.00). ISBN 0-939791-30-7


Ending Remediation: Linking ESL and Content in Higher Education. Sarah Benesch, editor. How higher education is bridging the gap between ESL and content courses through collaboration across the curriculum. Eight faculty members describe innovative linkage programs. 1988. $12.00 ($10.00). ISBN 0-939791-33-1

TESTING

Reviews of English Language Proficiency Tests. J. Charles Alderson, Karl J. Krahneke, and Charles W. Stansfield, editors. Descriptive and evaluative information on the major ESL/EFL tests used worldwide, including more than 40 reviews. 1987. $16.50 ($15.00). ISBN 0-939791-31-5


CLASSROOM PRACTICES

Classroom Practices in Adult ESL. Donna Ilyin and Thomas Tragardh, editors. Articles on topics such as classroom organization and management, evaluation, materials, cultural considerations, literacy and reading, communicative competence, and more. 1978. $8.00 ($5.00).

Classroom Practices in ESL and Bilingual Education. Muriel Saville-Troike, editor. A collection of articles designed to acquaint classroom teachers with successful methods and materials used in the ESOL and bilingual education contexts. 1979. $5.00 ($4.00).

ON TESOL SERIES

Selected papers from TESOL Annual Conventions contained in the following volumes:

On TESOL '84. Penny Larson, Elliot L. Judd, and Dorothy S. Messerschmitt, editors. 1984 in Houston.


On TESOL '82. Mark A. Clarke and Jean Handscombe, editors. 1982 in Honolulu.


Individual volumes available for $9.95 ($7.95).

Complete set of all 6 volumes of the On TESOL Series — $45.95 (35.95).
TESOL PAPERS


Guidelines for the Certification and Preparation of Teachers of English to Speakers of Other Languages in the United States. 1976.


All papers: single copies, $1.00 each; multiple copies in groups of 10 for $5.00.

TESOL CLASSICS ON SALE


The Human Factors in ESL. James E. Alatis and Ruth Crymes, editors. Collection of papers by ESOL experts: Fimocchiaro on the teacher, Paulston on communicative competence, Tucker on innovative approaches, Burt on error analysis, and more. 1977. Reg. $5.00 Sale $3.00.


New!

Resolutions Adopted by the TESOL Membership at Legislative Assemblies, 1971 to Present. 8½" x 11", 3-hole punched with covers, shrink-wrapped. $4.00 ($2.50).

TESOL NEWSLETTER SUPPLEMENTS


Writing and Composition. 9 articles on ESL writing and composition, plus tips for writing teachers and a brief list of books and journals. February, 1984. $2.50.

PUBLICATIONS AVAILABLE
Back Issues of the TESOL Quarterly

Identify your selection by volume and number:
TQ 10:2 means volume 10, number 2.

Numbers from 1968-1989 cost $8.00 each:
1968 Vol. 2, nos. 1, 2, 3, 4
1974 Vol. 8, nos. 1, 2, 3, 4
1976 Vol. 10, nos. 1, 2, 3, 4
1978 Vol. 12, nos. 1, 2, 3, 4
1979 Vol. 13, nos. 1, 2, 3, 4
1980 Vol. 14, nos. 1, 2, 3
1981 Vol. 15, nos. 1, 2, 3, 4
1982 Vol. 16, nos. 3, 4
1983 Vol. 17, nos. 1, 2, 3, 4
1984 Vol. 18, nos. 1, 2, 3
1985 Vol. 19, nos. 1, 2, 3, 4
1986 Vol. 20, nos. 2, 3, 4
1987 Vol. 21, nos. 1, 2, 3
1988 Vol. 22, nos. 1, 2, 3, 4
1989 Vol. 23, nos. 1, 2, 3, 4

Special price for all volumes listed above (2-23) $250.00
(Individually these volumes would cost $432.00)

Issues not listed above are no longer in print. They may
be obtained from:

University Microfilms International
300 North Zeeb Road
Ann Arbor, MI 48106

Back Issues of the TESOL Newsletter

Individual issues from 1978-1989 cost $3.00 each:
1978 Vol. 12, (5 issues) 1984 Vol. 18, (6 issues)

Volumes of the TESOL Quarterly and TESOL
Newsletter from the current year are available only
with membership in TESOL.
## TESOL ORDER FORM

**TEACHERS OF ENGLISH TO SPEAKERS OF OTHER LANGUAGES, INC.**

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Title</th>
<th>Unit Price</th>
<th>Total Price</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subtotal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Postage &amp; Handling</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

Mail your order to: TESOL Publications, 1600 Cameron Street, Suite 300, Alexandria, Virginia 22314-2705 U.S.A.

Credit card orders by fax: 703-836-7864, or telephone 703-836-0774.

**Prices**

Members pay the price in parentheses following the list price. **Orders not indicating a member number must pay nonmember price.** Your member number appears on your membership card and on the mailing label of every TESOL Quarterly and TESOL Newsletter that you receive.

All prices are subject to change without notice.

**Payment**

All orders must be prepaid unless accompanied by an official purchase order of more than $40.00 U.S. Checks (drawn on a U.S. bank) and money orders should be made out to TESOL. **You may use VISA or MasterCard with a $15.00 U.S. minimum purchase.** Please be sure to include card number, expiration date, telephone number, and cardholder signature.

**Orders from Outside the United States**

International Orders will be shipped foreign book rate. Please allow 3-4 months for delivery. To have orders shipped **airmail** please add 35% to the cost of the book, plus shipping charges.

**Claims**

Claims must be received within 3 months (6 months for claims from outside the U.S.).

**Shipping**

Postage and handling must be included in all prepaid orders, and will be added to credit card and invoice orders over $40.00 U.S. Shipping is to be figured using this scale.

<table>
<thead>
<tr>
<th>Order Subtotal</th>
<th>Shipping</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under $25.00 U.S.</td>
<td>$2.50 U.S.</td>
</tr>
<tr>
<td>$25.00 - $99.00 U.S.</td>
<td>10% of total</td>
</tr>
<tr>
<td>Over $100.00 U.S.</td>
<td>Declining percentages</td>
</tr>
</tbody>
</table>

Orders within the U.S. are shipped library rate. Please allow 4-6 weeks for delivery.

Ship to:

______

(Name)

______

(Address)

Telephone:

(Country Code) (Area Code) (Local #)

TESOL Member No. 0010000

Nonmember □

☐ Check enclosed ☐ VISA ☐ MasterCard

Card number: __________________________

Expiration date: _______________________

Cardholder’s signature: ____________________
Careers begin and grow in this book.


This eighth, revised edition of the Directory provides basic information about college and university programs in the United States leading to a degree or certificate in the teaching of English to speakers of other languages (TESOL).

Lists 158 institutions which offer 245 programs, including
- 30 programs for a doctor's degree
- 151 programs for a master's degree
- 26 programs for a bachelor's degree
- 31 programs for a certificate

This Directory is the first step toward identifying and answering individual needs.

ISBN 0-939791-35-8

Teachers of English to Speakers of Other Languages, Inc.
1600 Cameron Street, Suite 300
Alexandria, Virginia 22314 USA


I understand that this order must be prepaid with a check in U.S. funds made payable to TESOL or with appropriate credit card information. I will include postage & handling as indicated. (UPS for U.S. mailings only.)

Quantity: _________

Price: (check one):
☐ Member $15 _________
☐ Nonmember $20 _________

Postage & Handling: 15% _________ 20% _________ (UPS in U.S. only)

Total: _________

Check enclosed ☐ VISA ☐ MasterCard ☐

Card Number __________________________

Expiration Date ________________________

Signature _______________________________

Ship To: ________________________________

______________________________________

Send order and payment to TESOL, Inc.
1600 Cameron Street, Suite 300
Alexandria, Virginia 22314 USA
For the ESL Teacher
A Special Inexpensive Teaching Aid

The World Calendar
Le calendrier mondial
Der Weltkalender
El almanaque mundial
世界曆
الموافقون

The World Calendar is truly the first international teaching calendar, in six languages—English, French, German, Spanish, Arabic and Japanese. It is illustrated with photos which reflect the world's diversity of cultural tradition. This large (11"x14") calendar will be an attractive addition to any room—whether in the classroom or the office, and it will enrich each student's day with a wealth of international knowledge.

Provides:
• A helpful aid in teaching about other cultures. • An instruction item that can be used every day by the ESL teacher. • An easy to use cultural guide.

Includes:
• Holidays of the world's five major religions—Christianity, Judaism, Islam, Hindu, Buddhist—in six languages.
• Each weekday, month and photo caption in six languages.
• An International Time Zone Map.
• English radio broadcast frequencies throughout the world.
• An index to help you find, understand, and teach celebration days throughout the globe.
• The history and origin of calendars.

World Calendar – TESOL

Block Education Systems, P.O. Box 259, Clarks Summit, PA 18411

☐ Please send one (1) calendar at $9.95 + $2.75 postage ($12.70)*
☐ Please send ______ calendars at $9.95 + $2.00 postage ($11.95)
☐ Check or Purchase Order enclosed.
☐ Phone Orders 1-800-447-8561
☐ Charge my credit card. ☐ VISA ☐ MasterCard
Card No. ___________________________ Exp. Date __________

Name ________________________________

Address _____________________________________________________________________

City __________ State ________ Zip _______

*For airmail delivery outside the U.S. add $6.00
THE CALICO JOURNAL

The CALICO Journal is a fully referred, quarterly, professional journal that will keep you current on topics, problems, and concerns in this rapidly growing field. Membership in CALICO includes a subscription to the Journal and special announcements about CALICO programs and activities.

SATELLITE TELECONFERENCE

We continue to serve the profession with satellite teleconferences related to Using Technology in the Modern Language Classroom. Other CALICO Satellite teleconference broadcasts will follow.

SUMMER INSTITUTE

Our summer institute programs provide an excellent opportunity for participants to earn undergraduate/graduate university credit while enhancing their insight and skills in the area of applying computer technology. We continue to attract professionals from throughout the world.

ANNUAL SYMPOSIA

The annual CALICO Symposium provides a forum for discussing state-of-the-art technology in a variety of disciplines, with a view to adapting that technology to the more effective teaching and learning of languages. It includes intensive hands-on small-group seminars, papers, demonstrations, panels, and special interest group meetings.

CALICO SIGs

HyperMedia
CD-ROM
Computers in the Classrooms
Interactive Audio/Video
Foreign Character Fonts/Asian Languages
Courseware Development
Artificial Intelligence
Infowindow Systems

CALICO is an international consortium dedicated to applying high technology to the teaching and learning of languages. CALICO is composed of leaders and innovators representing academic, business, research, manufacturing, and government sectors who work together for the benefit of the language-teaching profession.

CALICO

3078 JKHB, BYU, Provo, UT 84602 (801) 378-7079 FAX: (801) 378-6533 BITNET: CALICO@BYUVAX
CAMBRIDGE APPLIED LINGUISTICS

LANGUAGE TRANSFER
Cross-Linguistic Influence in Language learning
Terrence Odlin
Odlin analyzes and interprets research, showing how much influence a learner's native language can have in making the acquisition of a second language easy or difficult. He provides a detailed look at work on other areas important for the study of transfer, including discourse, individual variation, and sociolinguistic factors.

SECOND LANGUAGE CLASSROOMS
Editor: Robert Keith Johnson
This authoritative collection of papers argues for the process of language curriculum development to be made explicit. The stages covered are: curriculum planning, specification of ends and means, program implementation, and classroom implementation.

THE LEARNER-CENTERED CURRICULUM
David Nunan
Grounded in the studies of what teachers actually do in their classes, this book develops the idea of a curriculum arrived at collaboratively between teachers and learners.

"...Should be examined carefully by all those involved in language curriculum work...refreshingly readable and eminently useable in a practical way"—Prospect

SECOND LANGUAGE CURRICULUM
Research on Teaching and Learning
Craig Chaudron
"...nobody who is seriously wanting to take account of research findings on second language classrooms can afford to ignore it"—WATESOL Newsletter

INTERACTIVE APPROACHES TO SECOND LANGUAGE READING
Editors: Patricia L. Carrell, Joanne Devine, David E. Eskey
This authoritative collection considers different models of reading as an interactive process and examines them from the perspectives of theory, research, and classroom applications.

LANGUAGE LEARNING AND DEAFNESS
Editor: Michael Strong
This collection of original papers draws upon work in linguistics, psychology, and education to highlight the relationship between language acquisition in deaf and hearing populations.

Cambridge University Press
Call toll-free: 800-872-7423, outside NY State: 800-227-0274, NY State only. MasterCard and VISA only.
Reviews of

ENGLISH LANGUAGE PROFICIENCY TESTS

Edited by J. Charles Alderson, Karl J. Krahmke,
and Charles W. Stansfield.

Reviews of English Language Proficiency Tests provides
descriptive and evaluative information on the major ESL/EFL tests
being used in the world today. By providing this information in a
consistent format, Reviews assists test users in selecting tests ap-
propriate to their needs and in evaluating the quality of those tests.

A description of the test and a discussion of its reliability, valid-
ity, and related issues are contained in the review. Also included
in each review are reference materials: references cited by the re-
viewer and, in many cases, references added by the editors to pro-
vide more comprehensive information on the test being re-
viewed.

Members $15.00  Nonmembers $16.50

Plus postage and handling.
All orders must be pre-paid.

TEACHERS OF ENGLISH TO
SPEAKERS OF OTHER LANGUAGES
1600 Cameron Street, Suite 300
Alexandria, Virginia 22314  USA
NEWBURY HOUSE

Excellence in Second Language Texts

Reading

Issues for Today
An Effective Reading Skills Text
Lorraine C. Smith and Nancy Nici Maior

Up-to-the-minute topics immerse intermediate learners in reading, discussion, composition, and outlining practice. Lesson-end quizzes, original photos, and a pull-out answer key enhance comprehension.

Expanding Reading Skills Advanced, 2/e
Linda Markstein and Louis Hirasawa

The revision of this successful text offers a process approach to reading, including pre-reading exercises, a solid writing connection, and an optional research paper.

Writing

Write Away
Books 1, 2, and 3
Donald K. H. Byrd and Gloria Gailingane

Sentence-level practice, contextualized grammar, and high-interest topics make the series a must for your low-level writing courses.

Listening

Noteworthy
Listening and Note-taking Skills
Phyllis L. Lim and William Smolzer

Based on the theme of the United States, this high-intermediate text helps students hone their listening and note-taking strategies.

Request your examination copies today by calling: 800-233-4190

NEWBURY HOUSE PUBLISHERS
A Division of Harper & Row
Cross Currents is a biannual publication of the Language Institute of Japan (LIOJ) which provides a forum for the interdisciplinary exchange of ideas within the areas of cross-cultural communication and language skill acquisition and instruction. Our goal at Cross Currents is to act as a medium through which international ESL/EFL issues, theories, and methodologies can be made available to our Japanese audience, and also through which Japanese ESL/EFL contributions—increasingly vibrant, innovative, and influential—can be disseminated to our international colleagues.

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

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

<table>
<thead>
<tr>
<th>SUBSCRIPTIONS</th>
<th>1 Year</th>
<th>2 Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual</td>
<td>US $15.00</td>
<td>US $25.00</td>
</tr>
<tr>
<td>Institutional</td>
<td>US $19.50</td>
<td>US $35.00</td>
</tr>
<tr>
<td>ISSN 0289-1239</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For air mail, include an additional US $6.00 for one year and US $12.00 for two years.

Subscriptions may be ordered through your local bookseller or directly from Cross Currents, LIOJ, 4-14-1 Shirayama, Odawara, Kanagawa 250, Japan. Checks must be payable to Cross Currents (LIOJ), in US funds, and drawn on a US bank. Back issues available through Cross Currents. 20% discount for orders of ten or more placed by individuals.

RETURN THIS AD WITH YOUR SUBSCRIPTION ORDER, AND RECEIVE THE CURRENT ISSUE AS A COMPLIMENTARY GIFT.
A New Edition Emerges...

Understanding and Using English Grammar
Second Edition
Betty Schrampfer Azar

- Improved organization.
- Expanded grammar points.
- New exercises.
- A more thorough appendix containing grammar review units and exercises.
- More lively illustrations.
- Now available in single and split editions (Volume A—verb forms, Volume B—complex structures).
- Workbooks and teacher's guides now available.

Discover 17 years of buried treasure from the TESOL Newsletter!

An invaluable teacher's notebook, this book will provide practical ideas to guide the novice and refresh the experienced teacher. It has complete indexes to landmark contributions to the profession. 272 pp.

Section 1. The State of the Art
Section 2. Language and Culture
Section 3. Standard English as a Second Dialect
Section 4. English for Special Purposes
Section 5. Teacher Training
Section 6. Other Areas of Professional Interest
Section 7. Language Assessment
Section 8. Sharing the Professional Wealth
Section 9. Language Learning: Methods and Models
Section 10. Teaching Communication
Section 11. Linguistics and Grammar
Section 12. Reading and Vocabulary
Section 13. Composition and Writing
Section 14. Classroom Materials and Techniques

ISBN 0-939791-29-3

YES! I need Selected Articles from the TESOL Newsletter as a resource for new ideas.

Please send me ________ copy(ies) at ________ $15 US (member) each

$16.50 US (nonmember) each

I have included postage & handling: ________ 15% book rate

________ 20% UPS (US only)

All orders must be prepaid. Total amount enclosed $_________ US.

☐ Check enclosed
☐ VISA
☐ Mastercard

Card No. ___________________________ Expiration Date ____________

Signature ________________________________________________

Ship to (please print)

(name) _________________________________________________

(address) ______________________________________________

_____________________________________________________________________

Mail to TESOL Publications, 1600 Cameron Street, Suite 300,
Alexandria, Virginia 22314 USA
TOEFL: Research Reports
Your Continuing Source of Vital Information

Initiated in 1977, TOEFL (Test of English as a Foreign Language) Research Reports have become important reading for professionals around the world who are involved with testing English as a foreign language.

We've made a considerable investment in researching key language testing issues and concerns. As the developer of TOEFL, ongoing research is an essential part of our mission to ensure that the TOEFL family of tests remains a valid and reliable measure of English proficiency.

To date, 32 studies have been completed, and 25 others are in progress. All of our studies are approved by the TOEFL Research Committee, a six-member panel of independent research specialists.

Recent titles include Context Bias in the Test of English as a Foreign Language and Confirmatory Factor Analysis of the TOEFL.

All reports are available at a modest cost. To get a FREE TOEFL Research Reports overview and order form, please fill out and return the coupon to the address below.

TOEFL Research Reports
Educational Testing Service
P. O. Box 6161
Princeton, NJ 08541-6161
USA

Please send me my FREE TOEFL Research Reports overview and order form.

Name ___________________________

Title ____________________________

Institution ________________________

Address __________________________

________________________________

Zip/Postal Code ____________________
SCENES
Saskatchewan Council for Educators of Non-English Speakers

presents/reçoit:

TESL Canada '91

Theme:
"PART OF THE PICTURE!"

Call for Papers deadline: Nov. 1, 1990

Dates: May 2-4, 1991

Place: University of Saskatchewan
      Saskatoon, Saskatchewan

For more information about the Call for Papers write:
SCENES
Box 340, Sub Post Office #6
University of Saskatchewan
Saskatoon, Saskatchewan
S7N 0W0

For general information about the Conference write:
SCENES Conference
c/o Conference Office
University of Saskatchewan
Saskatoon, Saskatchewan
S7N 0W0
NEW ESL TITLES FROM NCTE

Not Only English
Affirming America’s Multilingual Heritage

Harvey A. Daniels, editor

In this powerful collection, twelve outstanding scholars and educators provide the kinds of historical, linguistic, and educational backgrounds necessary to understand the language-restriction movement which is resulting in English-only legislation across the country. The book is divided into four sections: the first describes the nature, development, and extent of the contemporary English-only movement; the second looks at the potential impact of the proposed federal English Language Amendment; the third analyzes the causes and motivations of language protectionism; and the fourth suggests political and professional responses to the English-only movement. 135 pp. 1990.

No. 33639-1148  $8.95 ($6.95, NCTE member)

When They Don’t All Speak English
Integrating the ESL Student into the Regular Classroom

Patt Rigg and Virginia G. Allen, editors

This thorough and practical volume comes at a time when the need for guidance in teaching students who speak English as a second language is greatest. TESOL estimates that every classroom teacher will encounter at least one student who speaks English as another language. Many times these teachers do their best to integrate new speakers of English into the classroom, but they are relying on their own good sense—not on formal training. Rigg and Allen have compiled a collection of essays by notable figures in ESL teaching and research, who outline principles and techniques for working effectively with language-minority students in the classroom. One reviewer said, “It is by far the best thing I have read on English as a second language.” 176 pp. 1989.

No. 56932-1148  $12.95 ($9.95, NCTE member)

To order: Send check or money order to NCTE, 1111 Kenyon Rd., Urbana, IL 61801 (Include $1.35 p/h for the first book ordered and 65¢ for each additional.)

OR CHARGE BY PHONE—1-217-328-3870
(Minimum charge order, $15)
"We are writing to classroom teachers, both mainstream and ESL, to those who educate classroom teachers, and to those who study teachers and children working together in the classroom."

Edited by: Pat Rigg and D. Scott Enright

ESL Teachers as Language Advocates for Children...Courtney Cazden

ESL Children's Writing: What We've Learned, What We're Learning...Sarah Hudelson

Reading in ESL: Learning from Kids...Pat Rigg

A Children's Story...Carole Urzúa

Use Everything You Have to Teach English: Providing Useful Input to Young Language Learners...D. Scott Enright

$10.00 members, $12.50 non-members, plus postage.

Teachers of English to Speakers of Other Languages
Suite 205, 1118 22nd Street, NW, Washington, USA D.C. 20037
Power Tools
For People Who Build
Language Skills.

Once in a great while, new tools are created that are so powerful and versatile they do more than make a job easier; they transform the way it's done.

The new Franklin Spanish Master/Maestro Inglés is the world's first and only true bilingual electronic dictionary. But it can help a student in so many ways, it might better be called an electronic tutor. It offers complete translations for more than a quarter million Spanish and English words from the renowned Collins® Spanish/English dictionary. Plus, it gives instant access to grammatical information so students can use what they've learned. Conjugations, sample usages, correct word selection, personal word drills, tutorial games—all in both languages—come at the touch of a button. Students learn at their own pace.

The Franklin Language Master* LM-4000, the world's first speaking dictionary, is a powerful tool for expanding vocabularies. In addition to providing 274,000 definitions from Merriam-Webster's dictionary, it pronounces more than 83,000 words aloud. Students hear unfamiliar words, and use them confidently in conversation. The LM-4000 is also a thesaurus, with 487,000 synonyms on file. It corrects spelling, plays tutorial word games and will drill students on 3,300 common SAT words.

Complete teacher's manuals and curriculum materials are also available. To learn more, call our Education Line toll-free: 1-800-525-9673.

Franklin*
Learning Resources

Merriam-Webster® is a trademark of Merriam-Webster, Inc. © 1990 Franklin Learning Resources, Mr. Holly, NJ 08060
The EFL industry has produced thousands of books, hundreds of newsletters, scores of journals, dozens of magazines... and only one international newspaper

GAZETTE

gives you the inside news of what's happening in English language teaching.

Request your free sample copy from EFL Gazette,
10 Wrights Lane, London W8 6TA
Tel: (01) 938 4630 and find out what it's about.
Call for papers

International Conference on Teacher Education
in Second Language Teaching

April 17, 18, 19, 1991
Organized by Department of English,
City Polytechnic of Hong Kong

Deadline for proposals: November 30, 1990

The goals of the conference are to examine approaches to L2 teacher education, to discuss related research findings, and to examine options available in L2 teacher education, particularly those which focus on teacher development.

Proposals for papers and workshops are invited which address the following issues:
- developmental approaches in L2 teacher education
- research on L2 teacher education
- action research in L2 teacher education
- inquiry-based strategies in L2 teacher education
- innovations in teacher education practices

Proposals should be sent to:
Conference on Teacher Education in Second Language Teaching,
c/o Department of English,
City Polytechnic of Hong Kong,
83 Tat Chee Avenue,
Kowloon,
Hong Kong.

Fax: (852) 788 8894
Tel: (852) 788 8850
Current Perspectives On Pronunciation

MORLEY
BROWNE
CATFORD
CELCE-MURCIA
CRAWFORD
GILBERT
HUCKIN
TEMPERLEY
WONG

"Current Perspectives on Pronunciation will be welcomed by all teachers of English as a Second Language. Joan Morley has here assembled seven articles by authors who unequivocally believe that within a communicative framework there is the possibility—and even the necessity—for providing pronunciation instruction. These articles will satisfy all who insist on a sound theoretical basis for language instruction as well as those who are searching for practical methodology for the application of theory."

B.M. PETRIE
The Canadian Modern Language Review

Member $12, Nonmember $14, plus 15% postage & handling.

Teachers of English to Speakers of Other Languages, Inc.
1600 Cameron Street, Suite 300 • Alexandria, Virginia 22314 USA
Telephone 703-836-0774
Accent
English

Learning
the Sounds
of American
English for
Speakers of:
Spanish French
Chinese Arabic
Japanese Tagalog
Indonesian Korean
Vietnamese

Accent English is an audio-
cassette-based program that gives
precise auditory, visual, and oral
instructions on how to make the
sounds of American speech. The
program is currently available for
speakers of nine languages as
listed above. It targets those
sounds that a non-native speaker
of a specific first language may
have difficulty voicing.

The program is designed not
only to improve the user’s
pronunciation, but also to help
him or her acquire the appropriate
rhythm of the language, and the
proper stress, pause, phrasing, and
intonation patterns of American
English.

Designed for the advanced
learner, each program consists of
4 cassettes (90 min. ea.), a 144-p.
book, a mirror, and 42 visual-aid
cards that show a picture and
drawing designed to help the
learner shape his or her mouth to
produce the sounds of American
speech. Also included is a “Listen
to Me First” cassette with pre-
and post-test.

Each program is $79.50, plus
$4.50 shipping ($23.85 to non-
U.S.A. addresses). Be sure to
specify the version wanted for the
following native speakers: Arabic,
Chinese, French, Japanese,
Korean, Spanish, Vietnamese,
Indonesian, or Tagalog.

Full no-risk guarantee. Write
or Fax to place your order or
to obtain a free copy of our
ESL Catalog.

Audio-Forum
Room M927, 96 Broad Street
Guilford, CT 06437 U.S.A.
TEL: (203) 453-9794  FAX: (203) 453-9774
Summer Institute in English and Applied Linguistics

Language and Understanding

A two-week residential course for university or college lecturers, teacher trainers and senior teachers of English

14 – 27 July 1991

Director: Gillian Brown

The speakers will include:

Jean Aitcheson  Ellen Bialystok
Keith Brown     Lesley Milroy
Bernard Spolsky  Deidre Wilson

For details: Richard Mason, Madingley Hall, Madingley, Cambridge CB3 8AQ, England. Tel: (UK) 954 210636; Fax: (UK) 954 210677

Interested in Teaching in Japan?

Japan Network Services seeks interested ESL/EFL teachers for well paying positions in the Kansai (Osaka-Kyoto-Kobe) area of Japan. Jobs with universities, colleges, companies and language schools are open throughout the year. JNS provides relocation counseling while in the US and continues with personal service through agent in Japan. JNS has 8 years experience in the job market in Japan and endeavors to match jobseekers with desired openings.

Send resume and cover letter for more information to:
Japan Network Services
2049 Century Park East, 12th Fl.
Los Angeles, CA 90067
La buena vida
means directing your own prestigious language school while living in a lovely colonial style home minutes from the beach in Cancun, on the Mexican Caribbean.

Escuela Xicalango, established in 1976, is a private ESL academy for adults located one block from Cancun's main street. The eight well-equipped classrooms serve 400+ students daily year round.

The director's residence, with four bedrooms and three baths, is located on the same lot but has a separate entrance facing a quiet park. Investor's immigration status would allow the new owners to work in the school, assisted by fine staff members with years of experience.

For more information, please write Deborah Frisch, 650 Santa Fe Avenue, Albany, CA 94706
Don't get left behind

Get the EFL Careers Guide 1990/91

Better qualifications are the key to a better job in any profession - EF/SL is no exception. The EFL Careers Guide 1990/91 is your key to a better future in EF/SL. It will lead you through the maze of courses and qualifications available. From initial training to MAs, with the Guide's help you will be able to choose the very best option for your talents, needs, and pocket.

- initial training courses for 1990/91
- further qualifications for practising and experienced EF/SL professionals who want to progress - a clear overview of the bewildering range of options, including information on university and other courses in the US, Canada and Australia
- country-by-country details on what it's really like to live and work all over the world - because forewarned is forearmed!
- country-by-country directories of thousands of schools worldwide - invaluable when applying for jobs
- courses run especially for non-native teachers
- an introduction to specialist TEFL careers: ESP, English for Business, EFL marketing, administration, publishing and management

From reviews of the last Guide:
'I have to lock my copy up - I have found the Guide an excellent source of information for dealing with enquiries'
Kate Naameh, Marketing Director, International House, London

For your copy, send just £5.75 (UK) or £7.05/US$16.00 (overseas) (inc p+p) to: EFL Ltd (TQ), 64 Ormby Road, Ramsey, Isle of Man, UK. Please make all cheques payable to EFL Ltd.
NCTE RESEARCH FOUNDATION GRANTS AVAILABLE

The Research Foundation of the National Council of Teachers of English annually awards Grants-in-Aid for research that has significance for the teaching or learning of English/Language Arts or closely related fields.

Grants-in-Aid are awarded to professional researchers, including graduate students conducting dissertation research. The Foundation wishes to support the work of established researchers, as well as to encourage researchers just entering the profession. Most awards for grants range between $2,000 and $10,000. However, to encourage applications from established researchers undertaking larger studies, the Trustees will consider a very small number of proposals up to $25,000. Examples of recently funded studies:

- Symbolic Play as a Precursor to Children's Literate Behavior
- A Study of Student-Teacher Conferences in a Ninth-Grade Writing Class
- Course Notes in Rhetoric from 18th- and 19th-Century Scottish Universities
- Patterns of Discourse in Classroom Discussions of Literature
- Exploring the Use of Dramatic Activities in the Secondary English Classroom
- Learning to Read by Reading: Exploring Text Indices for Understanding the Process
- Pre-Service Education in Literacy-Related Areas: A Vygotskian Perspective

Deadline for proposals is February 15. For application forms or more information, write to Project Coordinator, NCTE Research Foundation, 1111 Kenyon Road, Urbana, IL 61801, or call (217) 328-3870.
ESL in higher education — remedial or credit-bearing?

Ending Remediation: Linking ESL and Content in Higher Education
edited by Sarah Benesch

The controversy over remedial versus credit-bearing ESL courses affects many colleges and universities. *Ending Remediation* will alert you to new possibilities for strengthening academic second language learning.

Eight ESL professionals suggest challenging, integrated curricula and approaches to campus-wide collaboration. Learn how to involve content and language faculties, native and nonnative students, administrators, and counselors.

— Required reading for anyone dealing with a postsecondary ESL program!

$10.00 members, $12.00 nonmembers plus 15% postage/handling. All orders must be prepaid. VISA, MasterCard welcome. 6" x 9" softbound. ISBN 0-939791-33-1

Teachers of English to Speakers of Other Languages
1600 Cameron Street, Suite 300
Alexandria, Virginia 22314 USA
703-836-0774
THE DYNAMIC INTERLANGUAGE
Empirical Studies in Second Language Variation
edited by Miriam R. Eisenstein

In The Dynamic Interlanguage, contributors investigate variation and factors contributing to variation in second-language acquisition, as papers examine the influence of context, relative social issues, target-language variety, and language-mediated studies. Theories of acquisition, analyses of interlanguage data, and suggestions for second-language pedagogy comprise sections focusing on • aspects of variation theory and how multiple variables interact • adjustment in native and nonnative discourse • roles played by alternative varieties in the acquisition process • additional factors—including surface-level constraints and social status—contributing to variation, and • pedagogical applications of what has been learned about second-language variation. Specifically, chapters explain why the consideration of multiple variable interaction must be built into theoretical models, how to choose a pedagogical norm for second-language teaching, and how some narrative differences may be incorporated into the ESL curriculum, among other topics. This volume will be of interest to teachers and administrators of ESL programs, language and linguistics professors and researchers, and educational psychologists. A volume in the series Topics in Language and Linguistics.

0-306-43174-2/342 pp./ill./1989/$49.50
($59.40 outside US & Canada)
text adoption price on orders of six or more copies: $32.50

PLENUM PUBLISHING CORPORATION
233 Spring Street
New York, NY 10013-1578
Telephone orders: 212-620-8200/1-800-221-9369

• OHIO UNIVERSITY •
1990 SUMMER COURSES
Linguistics Department • Ohio Program of Intensive English
International Training and Development

TESL INSTITUTE
June 11 - August 18

TEFL INSTITUTE
July 1 - August 24

Intensive professional development programs designed for elementary and secondary school teachers as well as trainers and supervisors in government ministries, international organizations, and universities. OPIE Institutes include:

• Experienced international ELT faculty
• Special courses of study for supervisors and trainers, experienced EFL instructors, and certified public school teachers seeking ESL validation
• Classes and workshops in methodology, curriculum design, evaluation and testing, observation, and teaching in difficult circumstances
• English refinement courses for professionals in EFL situations
• Field trips, social & cultural activities

Summer TESL/TEFL Institutes
Gordy Hall 201 TO Ohio University Athens, Ohio USA 45701
Phone (614) 593-4575 Telex 810-239-2992 Fax: 614-593-4577
TESOL MEMBERSHIP APPLICATION

Please check: □ New □ Renewal (ID No. 001-0000-__ __ __ __)

Name (Print) ________________________________

Mailing Address ________________________________

City ____________________________ Province/State ____________ Country ____________

Postal Code _______ Fax #: _______ Country Code _______ Area Code _______ Local #: _______

Telephone #: _______ Country Code _______ Area Code _______ Local #: _______

Please make check in U.S. funds drawn on a U.S. bank payable to TESOL or provide the appropriate credit card information and mail to: TESOL, Suite 300, 1600 Cameron Street, Alexandria VA 22314, U.S.A.

Check enclosed _______ VISA _______ MasterCard _______

Card number: ____________ Exp. date: ____________

Cardholder’s signature ________________________________________________________________

MEMBERSHIP CATEGORIES

Membership includes subscription to TESOL Quarterly and TESOL NEWSLETTER

Check appropriate box:

□ INDIVIDUAL ................................................................. $ 42.00

□ STUDENT (for those engaged in at least half-time study)* ........................................... $ 22.00

□ JOINT (two-member household) ................................................................. $ 60.00

□ INSTITUTIONAL/LIBRARY .............................................................................. $ 75.00

□ COMMERCIAL (for profit organizations) ................................................................. $250.00

*Student members are required to provide faculty signature, address, and telephone number to verify minimum half-time status. Participation is limited to three years.

Faculty Signature/Title ____________________________

Institution/Address ________________________________________________________________

Telephone ____________________________

ALL MEMBERS OUTSIDE THE U.S. must add one of the following postage fees to their membership dues:

(1) FOREIGN SURFACE MAIL—All countries outside the U.S. ........................................ $ 6.00

(2) NEWSLETTER BY AIR, QUARTERLY BY SURFACE ........................................ $15.00

(3) BOTH QUARTERLY AND NEWSLETTER BY AIR

Canada and Mexico ............................................................................................... $ 8.50

Caribbean, South America, Europe, and the Mediterranean ................................ $20.00

Africa, Asia, and the Pacific ................................................................................. $26.00

INTEREST SECTIONS (maximum 3 allowed)

MARK with (1) the Interest Section in which you wish to be active and vote.

MARK with (2) not more than two additional Interest Sections in which you have a secondary interest.

EI _______ English as a Foreign Language

FI _______ Intensive English Programs

EE _______ Elementary Education

SS _______ Secondary Schools

HE _______ Higher Education

BE _______ Bilingual Education

AE _______ Adult Education

AL _______ Applied Linguistics

RS _______ Research

RC _______ Refugee Concerns

TE _______ Teacher Education

CA _______ Computer-Assisted Language Learning

PA _______ Program Administration

MW _______ Materials Writers

DS _______ Teaching English to Deaf Students

VD _______ Video

AREAS OF WORK

____ Pre- and Elementary School, ______ College or University, ______ Secondary School,

____ Adult Education, ______ Other (Specify _______)

I wish to receive more information about TESOL:

____ Publications, ______ Conventions, ______ Awards and Grants, ______ Summer Institutes, ______ Affiliates,

____ Interest Sections, ______ Employment Information Service, ______ Insurance (U.S. addresses only)

376
TEACH OVERSEAS
with ELS International


Since 1977, over 1,200 ESL instructors have chosen ELS International for an overseas teaching experience. They have found teaching with ELSI to be personally and professionally rewarding. We attribute this both to what the teachers bring to their assignments and to what we provide by way of support:

- Established programs and curricula
- ELSI trained supervisors
- Year round openings
- Competitive salaries and benefits
- Furnished housing and airfare (some locations)

We invite you to write for more information on ELS International's overseas teaching opportunities. Please enclose a current resume with your letter.

ELS International Inc.
Recruitment Officer Number 4
5761 Buckingham Parkway, Culver City, CA 90230 USA
Tel: 1-800-468-8978 (outside California) 213-642-4618 (in California)
does it again

- Handfree automatic student monitoring.
- FOUR program distribution and Library or Free Study.
- Master program recorder includes selectable:
  
  Automatic repeat of single or multiple phrases,
  
  Edit and reformat tapes, adjust spacing for student responses
  and bookmarks for continuous tape segment replay,
  
  Automatic “end of phrase” pause, ideal for “answer only” oral
  testing or analyzer use.

- Student recorders include automatic last sentence repeat without
  head to tape contact reducing unnecessary wear on all components.

- Computerized design and circuitry result in simple operation and
  an effective teaching tool.

practical language lab systems
at affordable prices

Best of all the MC-9800 system with all of these features is still
competitively priced. Call us for a budgetary estimate which will
pleasantly surprise you. There is no obligation.

Toll Free (800) 553-7669

educational electronics corporation
213 North Cedar Ave • Inglewood, CA 90301

ECC/WICOM
BUILDING A REPUTATION . . . ONE LAB AT A TIME