LANGUAGE PROFICIENCY

Defining, Teaching, and Testing

Edited by
Louis A. Arena

Language Proficiency

Defining, Teaching, and Testing

Language Proficiency
Defining, Teaching, and Testing

Edited by
Louis A. Arena
University of Delaware
Newark, Delaware

Springer Science+Business Media, LLC

Library of Congress Cataloging-in-Publication Data

Language proficiency : defining, teaching, and testing / edited by
Louis A. Arena.
p. cm. -- (Topics in language and linguistics)
1. Language and languages--Study and teaching. 2. Language and
languages--Ability testing. 3. Applied linguistics. I. Arena,
Louis A. II. Series.
P51.L354 1990
418'.007--dc20 90-46926
CIP

ISBN 978-1-4899-0872-8 ISBN 978-1-4899-0870-4 (eBook)
DOI 10.1007/978-1-4899-0870-4

CONTENTS

TESTING FOR LANGUAGE PROFICIENCY

INTRODUCTION

Louis A. Arena

University of Delaware
Newark, DE

This monograph contains select, revised, and invited papers which deal with the topic, Language Proficiency: Defining, Teaching, and Testing. This topic was the theme of the eighth annual symposium held at the University of Delaware. The papers contained in this volume are invited papers or were originally scheduled for presentation and/or presented at the eighth annual Delaware Symposium on Language Studies. The papers combine research conducted in the areas of teaching, testing, and defining second language proficiency within the profession of applied linguistics. They are divided into three principal sections: "Applied Linguistics and Language Proficiency", "Language Proficiency in Reading and Writing", and "Testing for Language Proficiency".

In Part I, Paul Angelis' "Applied Linguistics: Realities and Projections re the Teaching Profession" sketches a historical portrait of Applied Linguistics, its definition, presence, and role in the profession that teaches second language proficiency. Angelis concludes that Applied Linguistics is still a young discipline in terms of substance, organization, and strategy, and that these three components will determine the prospects for the future of applied linguistics re the teaching profession. The next six papers address the issue of second language proficiency from various points of view. Kensaku Yoshida's essay "Knowing vs Believing vs Feeling: Studies on Japanese Bilinguals" concludes that some Japanese bilinguals are actually not necessarily bilingual because they very often face problems requiring other kinds of proficiency, i.e., cultural and sociopragmatic proficiency. Senko K. Maynard's paper "Understanding Interactive Competence in L1/L2 Contrastive Context: A Case of Back-Channel Behavior in Japanese and English" deals with interactive strategies across speech communities. In her study, Maynard articulates an important aspect of interactive competence, i.e., the management of conversation through back-channel strategies in casual conversation. The next paper, by Kyle Perkins and Sheila R. Brutten, applies several concepts from cognitive grammar, natural phonology, and item response theory in order to better understand how ESL students become proficient in processing the /D/ and /Z/ morphemes in English. They observe that the /D/ or /Z/ morphemes are stored as articulatory/perceptual images and are classified into prototypes whose allomorphs are known by their deviations from a prototype. Perkins and Brutten conclude that the assumption of unidimensionality in two data sets could not be absolutely shown because of various affective, cognitive, linguistic and test-taking factors. Robert Hammond's "The Affective Filter and Pronunciation Proficiency -

Attitudes and Variables in Second Language Acquisition" observes that some adult L2 learners make remarkable progress in acquiring L2 pronunciation, but then they experience a dramatic drop-off in their acquisition of L2 pronunciation. Hammond's study addresses the relationship between adult learner attitudes towards second languages and actual second language acquisition. Marion Lois Huffines' paper, "Proficiency without Support: My Parents Never Spoke It to Me", describes the results of a small study of Pennsylvania German speakers. One of her conclusions is that when Pennsylvania German converges towards English, the loss of the dative case in Pennsylvania German is an objective manifestation of the convergence between Pennsylvania German and English. Lilith M. Haynes concludes Part I with her paper, "Proficiency in Technical English: Lessons from the Bridge." In her delightful style, Haynes describes how students in the newer German university systems acquire proficiency in 'bridge courses' and what the lessons are that both students and teachers learn from Technical English courses.

Part II contains the results of several studies, theories, and methods for teaching second language proficiency in the skills of reading and writing. In "Training Formal Schemata--Replication Results", Patricia L. Carrell investigates the question of whether or not reading instruction which focuses upon the rhetorical organization of expository texts improves reading comprehension for poor reading-comprehenders. James F. Lee, in his paper, "Issues in Defining and Assessing Reading Proficiency: The ACTFL Guidelines", takes the different viewpoint that the proficiency level of readers should be characterized by how they interact with different types of texts, and not the structure of the text itself. The following papers in Part II deal with teaching proficiency in the skill of writing. In "What Composition Theory Offers the Writing Teacher", Patricia M. Dyer presents the need for composition teachers to recognize both the historical and the contemporary meanings of Composition Theory. Dyer stresses that both meanings inform the composition teacher, - the historical being durable and expandable, and the contemporary providing a forum for the exchange of teaching methods and research on writing, - and cites the significant contributions of various composition theories to the teaching of writing. Carl Mills' "Syntax and the Evaluation of College Writing: A Blind Alley" describes a study which once again addresses the relationship between written syntactic complexity and an increase in the quality of writing by freshmen university students. Mills provocatively concludes that syntactic complexity is not a strong predictor of writing effectiveness. Ghaida Salah also addresses the topic of syntactic complexity in her paper, "Discourse Analysis and Embedded Depth of Utterances: Clause Analysis Technique as a Measure of Complexity." Salah observes that embedding in discourse, both spoken and written, is related to proficiency of the second language learner and that the quality of discourse increases with increased embedding. James R. DeJong's paper, "The Role of Literature in Teaching ESL: Still Viable in the 21st Century," concludes Part II. DeJong's presentation contains cogent arguments for the ESL instructor to return to literature and cites several unique teaching applications that use literature as a basis, and which are just beginning to be touched on.

Part III addresses language proficiency from the viewpoint of testing. Marguerite G. MacDonald's "Communicative and Cognitive Language Proficiency: Implications for Testing in the L2 Classroom" addresses the timely issue of how and what is measured when testing communicative skills. MacDonald clearly distinguishes between cognitive ability and communication skills and shows why each must be considered separately in any evaluation of language proficiency. In his paper, "The Pragmatic Demands of Placement Testing," Francis J. Sullivan, Jr. examines how the overall ranking of university students' placement-test essays is influenced by the pragmatic form of the texts which the students produced. The results of Sullivan's study conclude that essay readers react positively to the communicative

function of text structure when rating essays for overall quality. He extends his conclusion to say that students' formal text structures take precedence over substantive communication in the eyes of the essay readers. Co-authors Kyle Perkins and Worthen N. Hunsaker's study focuses on the distinction between reading comprehension and information gain. They apply Bayes' Theorem to item analyses of cloze items in the paper, "A Comparison of Bayesian and Traditional Indices for Measuring Information Gain Sensitivity in a Cloze Test." They conclude that Bayesian indices are highly influenced by item difficulty and are unstable across samples; they guide the reader to some interesting directions for future research. Mary Ann G. Hood, on the topic of cloze testing, introduces another modification of a reading test in her paper, "The C-Test: A Viable Alternative to the Use of the Cloze Procedure in Testing?" Among several conclusions, Hood found that item difficulty also influences comprehension, but that even very difficult texts can be used for a C-Test. Her investigation of the C-Test raises several questions about text neutrality, scorability, and general reading skills which are worth further investigation. The topic of testing language proficiency is completed in Part III by S. Satyanath and T.S. Satyanath's paper, "Testing Language Proficiency in India." Not only are pedagogical issues discussed but also sociolinguistic issues, particularly three types of language variation, i.e., regional, social, and diglossic, are reviewed as to their influence on language testing in multilingual, pluricultural India.

The topic, Language Proficiency: Defining, Teaching, and Testing, is an important choice for a monograph. The papers collected here represent some of the leading research in assessing language proficiency while others suggest further steps towards a fuller understanding of second language teachers and learners. The importance of the topic re language proficiency is further underscored by many suggestions for further research and applications, contained in most of the papers included in this monograph. I think that this volume of select papers will stimulate even larger audiences of teachers, researchers, and test designers to guide the investigation of second-language proficiency well into the 21st century.

APPLIED LINGUISTICS AND LANGUAGE PROFICIENCY

APPLIED LINGUISTICS: REALITIES AND PROJECTIONS

RE THE TEACHING PROFESSION

Paul J. Angelis

Southern Illinois University
Carbondale, IL

The motivation for attempting a retrospective and projective view of applied linguistics with respect to language proficiency and the teaching profession, has, for me, been at least twofold. On the one hand, there is a professional perspective. By that I mean the actual work conducted by applied linguistics. What major and minor themes have been the focal points of research and general inquiry? What trends seem to have been developing in the directions for such work? At the same time, there is an organizational basis for such review. Here I refer to the formal establishment of associations and organizations. Why and how have applied linguists joined together in professional organizations? What prospects appear for the maintenance and growth of such organizations? The first of these issues touches on the seemingly age-old question of the nature and scope of applied linguistics. Admittedly, retrospection in this regard has often taken the form of introspection. Some would contend that constant attention to what activities are and can or should be encompassed within a field labelled "applied linguistics" detracts from participation in and dissemination of information on the work to be done. Such a commentary should not be dismissed lightly. There is value, nonetheless, in assessing the activities within applied linguistics because of the potential support that such reviews provide to the field, its participants, and its organizations. It can only provide needed recognition for work already accomplished and, by highlighting such activities, encourage more professionals to participate in a wider range of potential projects with a language base.

The structural and organizational issue is somewhat timely since this is the tenth anniversary of the American Association for Applied Linguistics (AAAL). Occasions such as this give added impetus to a forward and backward look at a discipline, especially one as recently established as applied linguistics. With this in mind, the primary focus of this review will be the professional and organizational status of American applied linguistics. It would be inappropriate as well as deceiving, however, to ignore the direction of applied linguistics within Great Britain, at least to the extent that there has been influence on American work. Actually, the influences have been mutual. Peter Strevens (1980) in his remarks offered upon the establishment of the American Association of Applied Linguistics in 1977, for example, points out the apparent and somewhat telling irony in the fact that, although the British Association of Applied Linguistics (BAAL) had been founded ten years earlier in 1967,

the decision to move in that direction as well as some prior actions such as the establishment of the School of Applied Linguistics at the University of Edinburgh were partly modelled after similar activities in the United States. He cites, in particular, the existence of the Center for Applied Linguistics in Washington, D.C. and its multidisciplinary approach to language issues. In recent years, the ties between the British and American Associations of Applied Linguistics have grown stronger. Since 1979 the journal Applied Linguistics has been jointly sponsored by both associations and 1988 will see the first jointly sponsored seminar, bringing together British and American applied linguists to share their thoughts on the status of "communicative competence" as a theoretical and practical construct within the field of applied linguistics.

Even within this American/British focus and our primary concern for fairly recent trends, it is helpful to establish some broader perspective on the nature and extent of applied linguistics. William Mackey (1968) writing for the British Council in Japan places the origin of the term "applied linguistics" in the United States in the 1940's. He credits the first uses of the term to "persons with an obvious desire to be identified as scientists rather than as humanists." Regardless of the motivation, there is little doubt that the timing and setting cited were in large part responsible for what did come to be known, at least in its early stages, as applied linguistics. Bertil Malmberg (1971) in his plenary address from the Second International Congress of Applied Linguistics in 1969 cites "essentially two reasons for the sudden birth of an applied linguistics during, and shortly after World War II: one theoretical and the other practical." The first he associates with changes in the orientation and methodology of linguistics from the 1920's and through the 1930's as a consequence of the advent of structuralism. The second was the war and post-war conditions which, particularly in the United States, created a need for a practical acquaintance with traditional as well as more exotic foreign languages. This clearly put the emphasis of early applied linguistics on matters related to language teaching. Given the stated circumstances, this was a natural development. While it is so often tempting to search for the shortcomings in early developments such as these, we must instead analyze the situation in proper perspective. Moreover, it is not overstating the case to attribute much of the credit for modern-day applied linguistics and all that it embraces on the foundation laid by the early work linked most directly to language teaching. William Moulton's summary (1961) of "Linguistics and Language Teaching in the United States 1940-1960" remains one of the most comprehensive treatments of that era.

In its formative stages a discipline can be evaluated not only by the nature of the work conducted under its name but also by the extent and manner in which the name has been adopted. In the case of journals and anthologies the term "applied linguistics" has been an attractive one. Whether it has been given primary or secondary billing, there has been a clear pattern of use on the part of authors, editors, and publishers to rely on "applied linguistics" as a term of reference. The journal Language Learning launched in 1948 continues to identify itself via its subtitle printed on every cover as A Journal of Applied Linguistics. The primary journal, which at its inception in 1962 chose to identify itself as the International Review of Applied Linguistics (IRAL), continues to do so but carries as well a further qualification in its full title as the International Review of Applied Linguistics in Language Teaching. Even journals with a more narrow focus such as educational technology have adopted the "applied linguistics" label. The journal System carries as its full title the International Journal of Educational Technology and Applied Linguistics.

Oddly enough, the link with language teaching has not always been associated with the formal use of the term "applied linguistics." Robert Lado's text on language teaching (1964), while clearly dealing with a linguistic orientation to the issue of how to approach language teaching situations, does not use the term. It is called, instead, Language Teaching: A Scientific Approach. Likewise, William Mackey's even more extensive survey of the same field published one year later (1965) is called simply Language Teaching Analysis. The closest, but still indirect reference, to the term comes in Halliday, McIntosh and Strevens' text of the same era (1964) entitled The Linguistic Sciences and Language Teaching.

A number of anthologies have appeared which have attempted to pull together a variety of information under the rubric of "applied linguistics." One of the earliest was Harold Allen's text entitled Readings in Applied English Linguistics. Appearing in 1958, the articles included were chosen to provide in a single volume a summary of the most significant ideas and issues which linguistics had then provided about the English language. The intended audience was teachers and prospective teachers of English. Although a small portion of the contributions dealt with English as a foreign language, the bulk of the text focuses on the native language situation. With articles on English linguistics, linguistic geography, usage, the teaching of composition, the dictionary, and the study of literature, the text served to fill a gap in English studies which for a long time had ignored the language itself. Once again, in a slightly different context, those who could survey the situation had sensed a place for merging the methods and results of linguistic inquiry with a language related field which could benefit from such a merger. Although somewhat less extensive and direct, such links are still found in the linguistic based publications, conference proceedings, and formal subgroups within such organizations as the National Council of Teachers of English and the Modern Language Association.

Other more recent anthologies have not had such a single focus. Instead, they have attempted to treat applied linguistics in a more comprehensive manner and to provide readers and users with a collection of articles dealing with ties between linguistics and a number of different areas of investigation. Ronald Wardaugh's Topics in Applied Linguistics (1974) is admittedly based on the language teaching situation. There is a noticeable movement to the foreign language context via sections dealing with second language teaching and contrastive linguistics but the sections dealing with spelling, reading, and language variation continue the treatment of linguistic based topics for foreign or native language contexts.

Just two years later (1976) Wardaugh and Brown published their A Survey of Applied Linguistics, which was a noticeable departure from previous such volumes in two respects. First, the coverage was much broader. Both first and second language teaching were included but also were topics such as bilingualism, dialectology, language and society, language disorders, and language testing. The second difference was perhaps more important than the first. While earlier volumes had been written to provide those who were not themselves linguists or applied linguists, especially language teachers, with information about linguistic based contributions, this text has a clear tone of dialogue with those who have already become initiated to the field. It was indeed a sign of the coming of age of applied linguistics.

A further indication of the maturation of applied linguistics at that time was the increased scrutiny given to the term itself, the rapidly expanding work being conducted under its name, and the role of applied linguistics in relation to other fields, especially linguistics. In the Wardaugh/Brown text the first chapter is devoted to the question "What is

Applied Linguistics?". The mid 1970's also saw a number of groups convened to address this same issue and to explore outlets for continued consolidation of the field. Out of these sessions came the publication of Kaplan's book (1980) entitled On the Scope of Applied Linguistics with contributions from British and American professionals. Organizationally, the need had at last been felt and barriers overcome to allow the formation of the American Association of Applied Linguistics.

It is difficult even today to point to any all-encompassing definition of applied linguistics. The efforts put into examinations of the field have, however, yielded some consensus on a number of basic principles. Among these four can be cited as central:

1. Applied linguistics can and should be a legitimate discipline in its own right.
2. The field is interdisciplinary in nature, drawing upon and contributing to linguistics as well as psychology, sociology, and a number of other disciplines.
3. The scope of the field must be open-ended. Specific issues should not be excluded in advance from consideration or investigation within applied linguistics.
4. The field should be problem oriented. In addition to any efforts to develop theoretical perspectives, a primary objective should be the solution of practical problems with a language component.

With slight variations, these themes have been strongly espoused and emphasized within American circles by Kaplan and Ferguson and from the British perspective by Strevens and Crystal. Kaplan's views have been summarized in the On the Scope of Applied Linguistics text and regularly in the volumes of the Annual Review of Applied Linguistics, the premier applied linguistics series introduced in 1980. Ferguson provides an added issue for consideration in the form of the special relationship of the field to linguistics. In his presentation on "Applications of Linguistics" at the 1974 Golden Anniversary Symposium of the Linguistic Society of America (1975) he commented on some of the perceptible shifts in attention within American linguistics. He cites as issues and challenges for the linguistic community three "social" problems in the form of national communication, mother tongue instruction and foreign language teaching and three "individual" problems -- deaf language, reading, and deception in language.

Strevens (1980) summarizes the British interpretation of applied linguistics as a "multidisciplinary approach to the solution of language based problems." His description of the rationale for the formation of the British Association of Applied Linguistics provides an interesting study of what he terms the "sociology" of such decision making. BAAL was and continues to be an independent organization with goals and activities related to but different from those of the Philological Society, the Linguistics Association, IATEFL, and the Modern Language Association. In identifying the applied linguists themselves, he describes them as typically "graduates, with some years of subsequent experience in some language-related field, who have then received advanced training in a range of disciplines which always includes linguistics and at least one other relevant subject, and who address themselves to the employment of the appropriate mixture of subjects according to the tasks they are engaged upon."

In one of the most detailed discussions of applied linguistics to date David Crystal (1981) properly distinguishes three closely linked aspects of the field. First, there is the issue of what is to be applied.

Unless we wish to argue for a complete abandonment of the term "applied linguistics" we must agree with Crystal that such a question will always be relevant and must regularly be the focus of attention for at least some applied linguists. This is by no means to say that applied linguistics is merely the application of linguistics. We have, presumably, moved beyond that misconception. It is, on the other hand. a recognition of the fact that linguistics remains a key element in all applied linguistics activity and that developments within linguistics should be relevant for applied linguists. The two other issues Crystal raises are, however, equally, if not more, important. There is the question of what represent the focal points of applied linguistic work and, lastly, the issue of how and why such activities can be pursued. On this last point Crystal provides an interesting commentary on how problems can be identified and discusses the question of how applied linguists can work with professionals from non-language based fields as so often they must. While his list is by no means exhaustive, Crystal's description of what he terms "first order studies" provides some framework upon which to project an agenda for applied linguistic activity. Within L1 problems he cites a whole range of issues dealing with spoken and written language acquisition and stylistic development. Within L2 problems there is a parallel range of issues for both teaching and learning as well as the domain of translation and interpretation. Where L1 and L2 situations coalesce there is in addition to areas dealing with bilingualism and multilingualism the matter of language planning and other contact phenomena.

Within such a backdrop of projections, claims, and summaries, it is appropriate that we examine the realities of the situation. What has actually characterized applied linguistic activity? Do the answers we derive indicate any trends? What procedural issues seem to emerge?

The first step in looking for work which may characterize a field is to begin at the broadest level. A fruitful source of data on completed works is the bibliographic listings of published and non-published material in references such as the ERIC collection. As a computerized network of entries across a broad range of fields, it is possible to see in list form what types of work have been conducted and, via the association of entries with descriptors, to assess to some degree the nature of work in a given field. "Applied linguistics" has been a term used as a major descriptor in the ERIC system since 1966. A computer search of published and non-published work in this field recently yielded 411 entries. These were almost evenly divided between 209 entries listed in Resources in Education (RIE) which are usually non-published items such as conference papers, occasional papers, proceedings, etc. and 202 entries in the Current Index to Journals in Education which are all published journal articles. Appendix A contains a listing of the topics included in the non-published work along with the number of entries for each. Topics with the same number of entries are not distinguished from each other in any way and the order of listing for those items has no significance.

The second entry shown labelled "general collections" contains a number of books such as the Wardaugh/Brown book cited earlier, collections from AILA and BAAL conferences and publications of the Center for Applied Linguistics. Except for these, however, the list contains a strong preponderance of entries dealing with language teaching. For those who may rely on such bibliographic compilations to gain a sense of the work being done in a given field, one wonders what impression this gives of applied linguistics. Two points should be noted, however. This collection covers the entire period from 1966 to 1987 and a good number of the language teaching entries come from the first half of that time span when applied linguistics was less active in other areas. Secondly, the ERIC system, especially for non-published work, depends heavily on both

designation of relevant material with ERIC and submission by authors. To the extent that either or both of these components of the system may fall short there will be gaps in which relevant work is not accounted for.

In the case of CIJE and the journal articles included the coverage is quite extensive. Appendix B lists the 76 journals included in the collection compiled under the "applied linguistics" descriptor along with the number of entries for each journal. Topics have not been included here but only the names of the journals. Except in the possible case of the IRAL entries listed first, there is a heavy emphasis again on language teaching topics with the largest number of journals being those dealing with the teaching of English as a first or second language and the teaching of French or German.

Obviously, a key element in this portion of the ERIC system is the designation of journals to be included. In this case a journal noticeably absent is Applied Linguistics. Through 1987 it has not been included in the system but it is to be added in 1988. For this reason and because of the special role of this journal in the United States and Great Britain we have examined articles which have appeared since its introduction in 1980. Appendix C provides that summary and shows a very different picture from the ERIC listing. A much broader spectrum of activity appears to characterize applied linguistics in recent years. Topics related to language teaching are included as well they should be but pragmatics, discourse analysis, lexicography, and cross-cultural issues are included as well.

A more timely indicator of work which applied linguists have shared with their colleagues is the pattern of presentations at AAAL annual meetings. Appendix D provides a summary of these presentations for the past five years, the last half of AAAL's ten year existence as a professional organization. Even more than in the case of the Applied Linguistics articles, the heterogeneous nature of these papers shows the expanding scope of activities in which applied linguists have been involved. A few points can be mentioned by comparing the AAAL papers with the Applied Linguistics articles. Areas such as second language acquisition and discourse analysis are prominent in both and indicate that the most significant amount of activity may be concentrated in that area. A few oddities do appear. One wonders why pragmatics, which has received a significant amount of attention in print (17 articles in Applied Linguistics) has had so little representation at AAAL meetings. The same is true to a lesser degree for lexicography.

From an opposite perspective, one wonders why topics such as bilingualism, language maintenance, and translation have not made their way into the pages of Applied Linguistics. Here, of course, such critical comments are not entirely justified since there are other outlets where publications on such topics can and have appeared. The relatively low ranking of sociolinguistics studies in Applied Linguistics is no doubt a case in point. This does raise an issue which must be considered in assessing future prospects for applied linguistics. For those both within and outside the field some confusion appears in trying to discover what work is taking place or even in trying to disseminate information on work completed. Should psycholinguistics and sociolinguistics be excused from applied linguistics because they are now firmly established as specialties in their own right? And in considering journals, what about the case, then, of one such as Applied Psycholinguistics? The situation is far from neat and clear-cut. But that is only to be expected in a field so recently established and so diffuse in nature. Moreover, its very diffuseness is a positive feature if the field is really to establish itself as a recognized discipline.

In terms of visibility and access to information, 1980 marked a significant turning point for applied linguistics. Not only was that the first year of publication for Applied Linguistics but it was that year which saw the introduction of perhaps the most useful tool available to applied linguists, the Annual Review of Applied Linguistics. A glance at the subject index for the seven volumes which have appeared to date shows an impressive array of topics ranging from adult literacy to bilingualism, language policies, and language in the professions. A young discipline requires nurturing and shaping. Here too the editors of ARAL should be commended for their strategic approach to that series. The choice to provide a cycle of thematic issues along with volumes including a general coverage of appplied linguistics gives users both intensive and extensive access to work in the field. Moreover, the detailed bibliographies and cross indexes provided greatly facilitate the task of following up on any of the topics included.

Having reached this far, we cannot avoid the next step of looking ahead to prospects for the future. Here, I would comment from three perspectives -- substance, organization and strategy. By substance I mean the actual work of applied linguistics. We have now in the United States established some momentum. American applied linguists are more actively involved in a more diverse range of language related programs and projects than has ever been the case to date. Given the increased opportunity for exchanging information on this work, such momentum should continue. The number of academic programs in applied linguistics at the MA and PhD levels has increased, promising to send forth larger numbers of new professionals who are aware of and committed to the field. These trends must be commended and encouraged.

From an organizational point of view, AAAL should be looked upon as the chief professional organization in our field. Membership is increasing, activities are expanding, and ties with other organizations are growing stronger. The association with the Linguistic Society of America continues to be of mutual benefit to both organizations and the linguistic base provided by LSA gives a relevant orientation to AAAL's structure and activity. A positive development has been the recent establishment of ties with BAAL, our British counterpart, including our forthcoming joint seminar. American applied linguists have always been well represented within our parent international organization, AILA. If any commentary can be made in this regard it is only that more U.S. applied linguists need to be brought into the fold at home. This year's list of AILA participants showed 67 persons with U.S. affiliations or addresses, only 35 of whom were members of AAAL.

Finally, the question of strategy. Apart from organizational issues and the matter of scope there is the question of how applied linguists reach out to the world at large, deal with those problems that require solution and establish contact with those who are their peers in other fields. In the spring 1983 issue of the newsletter of the British Association of Applied Linguistics David Crystal comments on a review by Roger Shuy of Crystal's text Directions in Applied Linguistics. The crux of the commentary is what the authors refer to as a "proactive" vs. a "reactive" perspective for the field. Shuy argued in his review for "an account of applied linguistics in which problems are central, i.e., the way to get things done with language, whether or not a negative problem intrudes itself." Crystal had argued for "an account which saw applied linguistics as primarily a problem solving subject, where the problems are encountered and initially defined by those professionals (language teachers, speech therapists, etc.) for whom language is a means of earning a living." Crystal explains his somewhat negative view by reflecting on the many situations in which he himself found a proactive stance to be counterproductive. He cites examples such as the occasions in which he

has tried to be positive with literary critics, to introduce them to the illumination which a linguistic analysis of a text can provide and has been greeted with a range of reactions from incomprehension to hostility. Crystal continues by citing similar situations in dealing with teachers, therapists, translators, lexicographers, etc.

There is no doubt a good deal of truth in what Crystal describes. Many applied linguists can think of similar situations in which they have participated. The problem comes, I believe, in thinking only of such group settings. Crystal's reactive approach has its place where and when problems have been encountered and the language aspects of those problems have at least been hinted at within the field in question. In other cases, making the effort to open a dialogue with individuals in other fields may represent the proactive stance most likely to bear fruit. We do run the risk of "crying in the desert" if the proper groundwork is not laid. Some fields have been more responsive to the possibility of collaborative work with applied linguists than have others. In the latter cases, a more long range view may be more appropriate. In any case, there is sufficient work to be done and challenges to be raised at both levels.

REFERENCES

Allen, Harold B. (1958). Readings in Applied English Linguistics. New York: Appleton-Century-Crofts.

Crystal, David. (1981). Directions in Applied Linguistics. London: Academic Press.

Crystal, David. (1983). "Action, reaction and Inaction in Applied Linguistics." BAAL Newsletter, 17, pp. 22-29.

Ferguson, Charles. (1975). "Applications of Linguistics." In Robert Austerlitz, ed., The Scope of American Linguistics. Lisse: The Peter De Ridder Press.

Halliday, M.A.K., A. McIntosh and Peter Strevens. (1964). The Linguistic Sciences and Language Teaching. London: Longmans.

Kaplan, Robert B. (1980). The Scope of Applied Linguistics. Rowley, Mass.: Newbury House.

Lado, Robert. (1964). Language Teaching. New York: McGraw-Hill.

Mackey, William F. (1965). Language Teaching Analysis. London: Longmans.

Mackey, William F. (1968). "Applied Linguistics: Its Meaning and Use." Journal of English Teaching, 1 (3), pp. 6-16.

Malmberg, Bertil. (1971). "Applications of Linguistics." In G.E. Perren and J.L.M. Trim. Applications of Linguistics. London: Cambridge University Press.

Moulton, William. (1961). "Linguistics and Language Teaching in the United States: 1940-60." In Trends in European and American Linguistics 1930-1960. Edited by Christine Mohrmann, Alf Sommerfelt and Joshua Whatmough. Utrecht, The Netherlands and Antwerp, Belgium: Spectrum.

Strevens, Peter. (1980). "Who Are Applied Linguists and What Do They Do? In Robert B. Kaplan. On the Scope of Applied Linguistics, pp. 28-36. Rowley, Mass.: Newbury House.

Wardaugh, Ronald. (1974). Topics in Applied Linguistics. Rowley, Mass.: Newbury House.

Wardaugh, Ronald. and H. Douglas Brown, eds. (1976). A Survey of Applied Linguistics. Ann Arbor: The University of Michigan Press.

Appendix A
Topics from Resources in Education (ERIC)
APPLIED LINGUISTICS
1966- 86 (209 entries)

Topic (category)	Number of Entries
General Language Teaching	37
General Collections (books, working papers, conference proceedings)	28
Native Language Teaching (usually English)	15
What is Linguistics	13
Reading	13
Linguistics and ESL	13
Teacher Training	13
Coherence and Composition	11
Teaching of Spanish	8
Contrastive Analysis	6
Teaching French	6
Teaching German	6
Psycholinguistics and Language Learning	5
Teaching Literature	3
Literacy	2
Language Policy/Planning	2
Sociolinguistics	2
Translation	2
Teaching Latin	2
Speech Research	2
Language Varieties	2
Language and Culture	2
Bilingualism	2
Computational Linguistics	2
Linguistics and Technology	1
Philippine Linguistics	1
Acquisition of Portuguese	1
Research Design	1
Nature of Language	1
Careers in Linguistics	1
Pragmatics	1
Lexicography	1
Phonetics	1
Japanese Verbs	1
Discourse Analysis	1
Artificial Languages	1

Appendix B
Journals Cited in CIJE (ERIC)
APPLIED LINGUISTICS
1966- 86 (202 entries)

Name of Journal	Number of Entries
International Review of Applied Linguistics	14
Revue de Phonetique Appliquée	11
Français dans Le Monde	10
Études de Linguistique Appliquée	10
Babel	10
Linguistik und Didaktik	8
Language Learning	7
ELT Journal	6
Elementary English	6
Modern Language Journal	5
English Language Teaching	5
Neusprachliche Mitteilungen	4
Langue Francaise	4
English Quarterly	4
TESOL Quarterly	4
Deutsche Sprache	4
Contact	4
French Review	4
Rassengna di Linguistica Applicata	3
Journal of Chinese Language Teachers Association	3
Canadian Modern Language Review	3
Deutsch als Fremdsprache	3
Neueren Sprachen	3
Reading Research Quarterly	3
Hispania	2
Reading Teacher	2
NALLD Journal	2
Englisch	2
Linguistic Reporter	2
College English	2
Praxis des Neusprachlichen Unterrichts	2
Languages	2
Glottodidactica	2
Foreign Language Annals	2
Unterrichtswissenschaft	2
Yelmo	2
Journal of Verbal Learning and Verbal Behavior	2
English Journal	2
Journal of Applied Linguistics (Greece)	1
ESPecialist	1
Issues in Applied Psycholinguistics	1
System	1
British Journal of Language Teaching	1
TESL Talk	1
American Speech	1
Language Arts	1
Journal of Reading Behavior	1
Theory into Practice	1
Humanist Educator	1
College Composition and Communication	1
Slavic and East European Journal	1
Modern Languages	1
Incorporated Linguist	1
PASAA	1
Educational Review	1

Cuadernos Hispanoamericanos	1
International Journal of Man-Machine Studies	1
Zeitschrift für Dialektologie und Linguistik	1
American Foreign Language Teacher	1
Deutschunterricht	1
RELC Journal	1
British Journal of Disorders of Communication	1
Revista de Filologia Espanola	1
Langues Modernes	1
Zeitsprache	1
Florida Foreign Language Reporter	1
English Record	1
Schulpraxis	1
Language	1
German Quarterly	1
International Reading Association Proceedings	1
Journal of Business English	1
Today's Education	1
Grade Teacher	1

Appendix C
Summary of Topics
APPLIED LINGUISTICS
Vols. 1 - 8 (1980-87)

Topic	Number of Articles
Second Language Acquisition	27
Pragmatics*	17
Discourse Analysis*	14
Cross-Cultural Issues*	12
Teaching-Learning Issues	11
Lexicography*	10
Communicative Competence*	8
Special Purpose Teaching/Learning	7
Comprehension*	6
International English	4
Language Testing	3
Classroom Instruction	3
Vocabulary	2
Reading	2
Curriculum Issues	2
Sociolinguistic Studies	2
Pronunciation	1
Translation	1

*Includes articles from special thematic issues.

Appendix D
Papers from AAAL Annual Meetings
1983- 87 (131 papers)

Topic	Number of Papers
Second Language Acquisition	22
Discourse Analysis	18
Language Planning/Policy*	11
Classroom Language Learning*	9
Bilingualism	8
ESP*	7
Second Language Teaching	7
Translation*	7
Language Maintenance and Revival*	7
Native American Languages*	6
Sociolinguistics*	6
Speech Disorders	5
Clinical Linguistics*	5
Language Strategies and Power*	5
Speech Processing	3
Literary Analysis	1
Orthography	1
Language and the Hearing Impaired	1
Pragmatics	1
Language in Advertising	1

*Includes papers from special symposia or panels.

KNOWING VS BEHAVING VS FEELING:

STUDIES ON JAPANESE BILINGUALS

Kensaku Yoshida

Sophia University
Tokyo, Japan

Georgetown University
Washington, DC

INTRODUCTION

In September, 1986, Prime Minister Yasuhiro Nakasone of Japan caused quite a furor when he said in a speech delivered to the members of his party that Americans had lower literacy and intelligence rates than the Japanese because of the heterogeneic composition of the American population. Granted that the content of his speech itself was of a seriously controversial nature, the sheer fact that Mr. Nakasone apologized just a few days after the initial speech is even more interesting from the point of view of cross-cultural communication. The Japanese, in a sense, are an "apologetic" people: we apologize when we enter someone's house or office (shitsureishimasu), we apologize for not being able to serve the kind of food that a guest might prefer (nanimogozaimasenga), we apologize for giving gifts which might not meet the high standards of the receiver (tsumaranaimonodesuga), and we apologize when we receive gifts (wazawazasumimasen, moushiwakearimasen). Apologizing, to the Japanese, is a form of etiquette which pervades throughout the fabric of Japanese social life. It is expressed by the various "polite" or honorific expressions which comprise an important part of the Japanese language.

Abiding by the rules of the Japanese honorific system is both natural and essential for anyone living in Japan. However, this Japanese way of perceiving human relations and behaving is so deeply rooted in the Japanese mentality that it cannot easily be discarded even when he is in a foreign culture and speaking a different language. In other words, even when a bilingual Japanese speaks English in the United States, he is still hampered at times by the persistent cultural tendency to follow the Japanese form of social etiquette. I, myself, am a bilingual (at least, there is no single definition of bilingualism which might exclude me), and I have lived in the United States and Canada, off and on, for over ten years. However, there are still areas of communication where I have problems. For example, I hesitate to call anyone who is either older or of a higher status than myself by his/her first name. I feel much more comfortable calling many of my American colleagues, both in Japan and in the United States, by their title and family name. In other words, to borrow Brown and Gilman's (1960) terminology, the power dimension of human relations seems to be working stronger in me than the solidarity dimension. I also feel uneasy asking questions before a lecturer has finished speaking, and I am very poor at coming out with an outright "no."

From purely linguistic competence standards, I should be quite competent in conducting both my academic and daily affairs in English in an American environment. However, there have been times when I have doubted my proficiency in the use of English. Furthermore, even from a communicative competence perspective, although I might have knowledge of what to say with whom in what circumstances, that does not necessarily mean that I am able to perform accordingly. Moreover, even if I could perform in an "American" way if I consciously strived to do so, that does not mean that I feel comfortable doing so.

Culture had long been considered an important aspect of second language acquisition (cf. Lado, 1957; 1964). However, with the emphasis on more universal trends in linguistic analysis (Chomsky, 1965), differences between languages as well as cultures were somehow pushed into the background. In recent years, however, they have re-emerged in discussions on cross-cultural pragmatics (Tannen, 1984) and in defining communicative competence (Saville-Troike, 1982).

The general impression we get from studies emphasizing the importance of culture seems to be that, for a foreign language learner to become really proficient in the target language, he/she must acquire not only the language but must also become acculturated psychologically and sociologically to the people and the culture of that speech community (cf. Schumann's (1978) acculturation hypothesis). Although this might essentially be true, there are studies which suggest otherwise. Lukmani's study (1972) showed that instrumental motivation is a very strong factor in the development of English proficiency in countries where English is used for official purposes, and Kiyono (1986) has shown that, in a foreign language environment like Japan, it is not whether a student wants to become a member of the target language community, but whether he/she is integratively motivated to learn about the target people and culture that correlates positively with their linguistic ability.

The purpose of this paper is to show that even second language learners, like Japanese bilinguals, who have lived in the United States and learned English in a second language environment do not necessarily think, behave, or feel like Americans. Bilinguals are not necessarily people who are bicultural to the extent that they behave and feel like a monolingual speaker of the languages they have acquired. In fact, in certain situations, a bilingual may not be able to perform "proficiently" in either language. Thomas (1983) notes the difficulty of teaching sociopragmatic rules to foreign language learners. Our contention is that some sociopragmatic rules are difficult even for second language learners.

1. The unique conceptual field of Japanese bilinguals

Ervin (1968) conducted a study in which she showed that Japanese wives of American soldiers who had come to the United States after World War II reacted differently to verbal stimuli presented to them in Japanese and English, even though the dictionary meanings of the stimuli were the same. For example, she showed that although the subjects would associate the word "New Year's Day" with words like "party" and "holiday", the same subjects when given the same stimulus in Japanese, "shougatsu." would respond with words like "omochi (rice cake)" and "kimono." New Year's Day is, of course, a very significant day in Japan from the cultural perspective. On the basis of her findings, Ervin concluded that co-ordinate bilinguals--bilinguals who have acquired two languages in different environments--utilized different conceptual fields depending on the language they were using at the time. In other words, they would be operating in a typically Japanese conceptual field when they were speaking Japanese, and in an English conceptual field when they were speaking English.

Kolers (1963) had further shown on the basis of word association tests conducted on Spanish-English, Thai-English, and German-English bilinguals, that although the bilingual's word associations differed between languages the majority of the time, there were still cases where the responses were semantically the same. He concluded from his observations that whereas concrete words triggered similar responses in both languages (e.g., table, lamb), abstract words tended to trigger culturally-specific responses (e.g., freedom, jealously).

The results of these studies suggest that bilinguals will respond differently to verbal stimuli given in the two languages they know if the stimulus is an abstract word, or, as in the case of Ervin's experiment, has an implicit cultural meaning behind it.

Although these studies, as well as the one by Wakabayashi (1973) in which the author compared word association differences between Japanese and American monolinguals, are interesting in themselves, they seem to imply that once you have acquired a certain language, you have also acquired all the cultural and conceptual connotations that go along with it. However, that does not seem to be so. As we speak of the concept of "interlanguage," our contention is that most bilinguals possess different versions of what one might call "interculture."

For our experiment, we selected 35 college students who had entered Sophia University via the special Returnee Entrance Examinations which Sophia administers to those students who have lived abroad for at least two years prior to entering college. All of them had, at one time or other, lived in the United States and had attended American Schools. As control groups, we selected 32 monolingual speakers of Japanese from the Japanese literature department of Sophia University, and 21 students at the University of Pennsylvania acted as our American control group.

The first task was a word association test. The bilingual group was asked to respond in Japanese to Japanese stimuli, and in English to English stimuli. The Japanese and English stimuli were given in a different order and administered one week apart. The control groups were asked to respond in their respective native languages.

Following Ervin's (Ervin-Tripp) experiment and taking into consideration the fact that different types of words evoke different types of associations, we selected words from the following categories for our stimuli: 1) nature (haru/spring, aki/autumn, tsuki/moon, hana/flower), 2) daily life (otoko/man, onna/woman, sensei/teacher, 3) society and ideas (jiyuu/freedom, seifu/government, nihon/Japan, amerika/America), 4) culture (kurisumasu/Christmas, shougatsu/New Year's Day).

In analyzing our results we 1) compared the percentages of responses for each stimulus word, 2) grouped the responses into semantic categories (see below), and 3) calculated the degree of agreement between the bilingual group and the respective monolingual control groups.

The semantic categories employed were the following: 1) context (CON)-- the response word is commonly seen in the context of the meaning conveyed by the stimulus word (e.g. SPRING → flower); 2) equal rank (ER) the response and the stimulus words are of the same taxonomic level (e.g. SPRING → autumn); 3) higher rank (HR)--the response word is from a higher taxonomic level than the stimulus word (e.g. SPRING → season); 4) lower rank (LR)--the response word is from a lower taxonomic level than the stimulus word (e.g. SEASON → spring); 5) part (PRT)--the response word constitutes a part of the object signified by the stimulus

word (e.g. TREE → leaf); 6) attribute (ATT) -- the response word shows a certain characteristic of the stimulus word (e.g. SPRING → warm); 7) symbolic (SYM) -- the response word is normally used as a symbol of the object signified by the stimulus word (e.g. AMERICA → stars and stripes); 8) phonetic association (PA)-- the response word is related sequentially to the stimulus as a part of the same idiom or well-known phrase (e.g. AMERICA → the beautiful); 9) personal (PER)-- the response word can apparently be attributed to the respondent's personal experience (e.g. TEACHER → Mr. Jones); 10) others (OT)-- the response word cannot be categorized in any of the above categories (e.g. FLOWER → spaceship).

The results of the responses in the nature group can be seen in tables 1a to 4b. The bilingual group's responses in both Japanese and English were quite similar to that of the Japanese control group, while differing more greatly from that of the American control group. However, the word "hana (flower)" did evoke different responses between the bilingual and Japanese control groups, although the responses of the bilinguals themselves were similar in both English and Japanese. It is interesting to note that, even though the flower most commonly associated with Japan and the Japanese is "sakura" (cherry blossom), to the extent that the word "hana" appearing alone often means "sakura," the bilingual subjects associated both "hana" and "flower" with the word "rose/bara," which has a definite western connotation. Yet, "rose" was not seen in the responses of the American subjects.

When we look at the words in the daily life category (tables 5a to 7b), we notice that although all three groups show a similar response pattern for the stimuli "man/otoko" and "woman/onna," the American control group has a lot stronger tendency to give a common response. One interesting phenomenon which was observed with the word "teacher (sensei)" was that there were individual responses by the bilinguals which did not appear at all in the Japanese control group (shinrai (trust), yuujou (friendship), sonkei (respect)).

The society and ideas category is similar to Kolers' abstract words. As in Kolers' study, our analysis also showed that the bilingual group responded quite differently depending on which language they were using. From tables 8a to 11b it can be seen, for instance, that the bilingual subjects responded to the English stimulus word "freedom" with such words as responsibility, myself, human beings, and independence which did not appear at all with the Japanese translation "jiyuu." Also, in responding to the stimulus word "government" and its Japanese equivalent "seifu," the bilingual subjects tended to associate mostly Japanese political events with the Japanese stimulus, although "gikai" appears instead of "kokkai," the Japanese parliament, or diet; whereas the English responses included some references to American politics as well. The bilinguals' English responses to the word "Japan" is also interesting in that words such as "Fuji" and "Asia" which would normally be used by foreigners viewing Japan from the outside are used, whereas, in Japanese, these associations do not occur. In fact, they resemble to an extent the American subjects' responses "Tokyo" and "China," the outsider's view.

Finally, the responses to the words in the culture category did, as we expected, reveal quite a difference depending on the language used. This was especially true in the case of "New Year's Day" versus "shougatsu."

From these results alone, however, it is difficult to determine whether the bilingual group is either the same or different from the Japanese control group.

In order to clarify this point, we took the three most popular responses

for each item in the Japanese and American control groups, calculated what percentage of the responses they accounted for, and then performed a t-test to determine how similar or different the bilingual group was to the control groups. We first found that in the American control group, the top three responses accounted for 54% of all the responses, whereas, in the Japanese control group they accounted for 39%. The fact that Americans tend to show more unified responses in word association tests than other peoples had already been shown by Wakabayashi (1973) and Rosenzweig (1961), and our results confirmed their findings.

Next, when we matched the top three responses of the respective control groups with the responses of the bilingual group, we found that 31% of the bilingual group's Japanese responses matched that of the Japanese control group, and 21% of its English responses matched that of the American control group. Furthermore, t-test results showed that the bilingual group's Japanese responses were significantly different from those of the Japanese control group ($t=2.39$, $p < 0.025$), and their English responses were even more significantly different from those of the American control group ($t=9.83$, $p < 0.001$). In other words, the Japanese bilinguals formed a totally different group from either the monolingual Japanese or American groups. The experiment showed that when a bilingual Japanese uses either Japanese or English, there is a strong possibility that he/she will be using a different conceptual field than the monolingual speakers of either language.

However, granted that the conceptual field of a bilingual is somewhat different from that of a monolingual, this only infers that bilinguals have a conceptual basis for coming up with different thought patterns from monolingual speakers. It does not necessarily mean that they will behave or feel differently on the affective plane as well.

2. Perceived social distance of Japanese bilinguals

In order to find out how our bilingual subjects felt (affectively) towards the Japanese and American people and their way of life, we took Acton's (1979) Perceived Social Distance test and measured the affective distances the subjects showed towards the two peoples. The Perceived Social Distance test is based on the semantic differential principle, although, instead of the seven-point scale, a six-point scale is used. It asks the subject three questions per item: 1) how does the subject himself feel about a certain item, 2) what does the subject think his own countrymen feel about it, and 3) what does the subject think Americans feel about it. For example, a subject would be given the item FUTURE and be asked to mark on a scale ranging from "predictable" to "unpredictable," how the subject felt, what he thought his own countrymen felt, and what he thought Americans felt.

The results of the Perceived Social Distance test were correlated with the word association scores to see whether a person whose conceptual field was closer to that of an American would also feel closer to Americans. The correlations showed that, contrary to the acculturation hypothesis, there was a significant negative correlation between the two scores when the difference in the distance perceived towards the Japanese and that towards the Americans was correlated with the subject's association score ($r= -0.457$, $p < 0.05$). In other words, the more a bilingual's conceptual field came to resemble that of an American, the farther he would feel himself to be from him affectively.

Strange as this may seem, similar results can be seen in other studies on Japanese returnees as well (cf. Iwao & Okamoto, 1984). As Savignon (1983) notes, the more a L2 learner acquires the target language and becomes capable of participating in the target culture, the more he might

discover things which he does not agree with, thus setting the stage for the development of a new interculture identity -- different from both his L1 and L2 cultures. (Cf. also Brown's (1980) hypothesis of the cultural critical period.)

From these results, it seems possible to conclude that Japanese bilinguals, at least, do not necessarily perceive the world like Americans, nor do they feel psychologically close to them; thus, creating the kind of problems in communication mentioned at the beginning of this paper. However, a problem remains. Is this true of every Japanese bilingual? After all, what we have seen is merely a group tendency, not a true picture of the individual.

3. The effects of age and length of stay

In this final section, we will take a brief look at studies conducted on Japanese bilinguals in which the factors of age and length of stay in the United States were seen to play very important roles in how a bilingual behaved or felt as a member of monolingual communities whose languages he knew.

In her study of 114 Japanese children living in the United States, Minoura (1984) observed that cultural assimilation consisted of three levels: 1) cognitive, 2) behavioral, and 3) affective. She interviewed each child, analyzed the responses, and grouped the subjects into the following 5 assimilation types: 1) subject does not recognize the difference of behavioral patterns between Japanese and Americans; 2) subject recognizes the differences but cannot (does not) behave like Americans; 3) subject is Americanized on the cognitive and behavioral levels, but not yet at the level of affection; 4) subject is Americanized on the cognitive and behavioral levels, but is still affectively unsettled; and 5) subject is Americanized on all three levels.

Minoura's results showed that the age of leaving Japan (9 years old) and length of stay in the United States (4 years) were significant factors in dividing the subjects into the different assimilation types (cf. table 14). She found that over 90% of the subjects who had left Japan under the age of 9 and had been in the United States for over 4 years belonged to either type 4 or 5. On the other hand, subjects who had left Japan after age 14 and lived in the United States for less than 4 years had, at best, only reached the level of behavioral assimilation, but not affective assimilation. Furthermore, Minoura suggests that the age range between 9 and 11 is a crucial period in which a child's cultural identity is determined (note that there are no type 4, i.e., affectively unsettled subjects, after the 9-11 age group).

To make Minoura's theory more concrete, let us take the following example. When asked the following 3 questions, "Do you know that Americans are frank?" (cognitive), "Can you be as frank as an American?" (behavioral), and, "Do you feel comfortable being frank?" (affective). Type 1 subjects would not even know that Americans were frank, type 2 subjects would know, but not be able themselves to be as frank, type 3 subjects could be frank if they tried, but would not feel comfortable, type 4 subjects could be frank, but be unsettled in their feelings, and type 5 subjects would feel comfortable being frank.

Although Minoura does not compare the bilinguals' responses with those of either monolingual Japanese or Americans, and, therefore, we do not know whether even type 5 subjects are really Americans inside or not, the fact that differences in age of entering and length of stay in the United States influence the level of assimilation of the bilinguals to American culture is significant.

In another study on Japanese bilinguals, Nemoto (1986) studied the bilinguals' use of Japanese expressions in typical Japanese situations requiring polite and indirect forms of expressions (giving gifts, responding to praise, and objecting to someone's opinion). She designated 9 different addressees (younger brother/sister, elder brother/sister, father/mother, close friend, classmate, under-classman, upper-classman, professor, and guest) to whom the expressions were to be addressed. Nemoto also compared the bilinguals' responses with those of monolingual Japanese as well as monolingual Americans who were given an English translation of the test. Although her study concentrated on the use of Japanese and not English, her results showed that age of arrival in the United States, length of stay there, as well as age of returning to Japan all influenced the assimilation level of the bilingual into Japanese culture. Her results are interesting because they reveal how similar or different the bilinguals are in comparison with the monolingual Japanese as well as Americans by showing that depending on the three factors of age and length of stay and their combination, a bilingual would be able to use polite and indirect expressions towards different numbers of addressees from the monolingual Japanese and Americans, thus confirming our hypothesis that bilinguals are at different stages of interculture, and that age and length of stay are factors which influence their sociopragmatic behavior and feeling.

4. Conclusion

In this paper I have tried to show that Japanese bilinguals are not necessarily bicultural, in the sense that even if their linguistic proficiency was quite high, they would very often face problems in situations which required proficiency in dealing with cultural and sociopragmatic matters. If this is so even in a L2 environment, how much more or less can we expect from our students in a foreign language environment? If we were to incorporate Minoura and Nemoto's assimilation levels, we would probably have to say that, since most Japanese begin learning English from the age of 13, they will never really be able to feel like an American, and that, some may never even be able to behave in an American way. The most we can expect is that they will be able to understand and acknowledge the differences which exist between Japanese and Americans. We might, therefore, have to conclude that, for the foreign language learner, we can only expect proficiency in the use of language at the pragmalinguistic level, and not at the sociopragmatic level, as was noted by Thomas. Furthermore, we have seen that even in the case of L2 learners, complete assimilation or acculturation is not always a realistic goal. The results which indicated that the more a bilingual's cognitive field comes to resemble that of Americans, the more he/she will feel distanced from them, suggests that, unlike the concept of interlanguage which posits the native linguistic competence as the final ideal state, interculture need not posit the native speaker's cultural concepts as its ultimate goal -- at least on the affective level.

I do not know how much interculture should be, or can be incorporated in defining proficiency. However, whether this notion is included or not, it must not be forgotten that even if a person has near-native linguistic proficiency, that does not mean that he/she has near-native sociopragmatic or cultural proficiency as well.

Note: The word association and perceived social distance experiments were supported by the Interdepartmental Research Fund of Sophia University. Thanks are due to Yo Matsumoto, Chiyoko Nemoto, Junko Hibiya and the members of the Sophia Applied Linguistic Circle for their cooperation in conducting the experiment as well as analyzing the data.

BIBLIOGRAPHY

Acton, William R. Perception of Lexical Connotation: Professed Attitude and Socio-Cultural Distance in Second Language Learning, Ph.D. dissertation, University of Michigan, 1979.

Brown, H. Douglas. Principles of Language Learning and Teaching, Englewood Cliffs: Prentice-Hall, 1980.

Brown, Roger & Albert Gilman. "The Pronouns of Power and Solidarity," in Styles in Language, ed. Sebeok, Cambridge, Mass.: MIT, 1960.

Chomsky, Noam. Aspects of the Theory of Syntax, Cambridge, Mass.: MIT,1965.

Ervin, Susan M. "An Analysis of the Interaction of Language, Topic, and Listener," in Readings in the Sociology of Language, ed. Fishman, The Hague: Mouton, 1968.

Iwao, Sumiko & Kouichi Okamoto. "Kokusai kankakuno jisshouteki rikaieno yobiteki kenkyuu," Keio Gijuku Daigaku Shinbun Kenkyuusho Nenpou, 1984.

Kiyono, Reiko. Attitudes and Motivation in Learning English -- in Case of Japanese High School Students, M.A. thesis, Sophia University, 1986.

Kolers, P.A. "Interlingual Word Associations," Journal of Verbal Learning and Verbal Behavior, 1963.

Lado, Robert. Linguistics Across Cultures, Ann Arbor: University of Michigan, 1957.

-------. Language Teaching, New York: McGraw-Hill, 1964.

Lukmani, Yasmeen. "Motivation to Learn and Language Proficiency," Language Learning 22: 261-273, 1972.

Minoura, Yasuko. Kodomono Ibunka Taiken: Jinkakukeiseikateino Shinrijinruigakuteki Kenkyuu, Tokyo: Shisakusha, 1984.

Nemoto, Chiyoko. Assimilation of the Japanese Returnees from the United States: a Sociopragmatic Study. M.A. thesis, Sophia University, 1986.

Rosenzweig, M.R. "Comparisons among Word Association Responses in English, French, German, and Italian," American Journal of Psychology. 1961.

Savignon, Sandra J. Communicative Competence: Theory and Classroom Practice, Reading, Mass.: Addison-Wesley, 1983.

Saville-Troike, Muriel. The Ethnography of Communication: An Introduction, Oxford: Basil Blackwell, 1982.

Schumann, John. The Pidginization Hypothesis: A Model for Second Language Acquisition, Rowley, Mass.: Newbury House, 1978.

Tannen, Deborah. "The Pragmatics of Cross-Cultural Communication," Applied Linguistics 5: 189-195 (1984)

Thomas, Jenny. "Cross-Cultural Pragmatic Failure," Applied Linguistics 4: 91-112, 1983.

Wakabayashi, Nobuchika. "Nihongono imirensouno zureto Eigokyouiku," Eigokyouiku, 1973.

Yoshida, Kensaku. " Kikoki-shijono gengo-shinrigakuteki tokushitsu:tango rensouto Social Distance no chousakara" in Kikoki-shijono Doukou Chousano Houkokusho, Tokyo: Sophia University, 1985.

Universal Legend
for Tables 1-13b

Legend
JM = Japanese Monolingual
BJ = Bilingual - Japanese
BE = Bilingual - English
AM = American

CON = Context

ER = Equal Rank
HR = Higher Rank
LR = Lower Rank
PRT = Part
ATT = Attribute
SYM = Symbolic
PA = Phonetic Association
PER = Personal
OT = Others

(%)

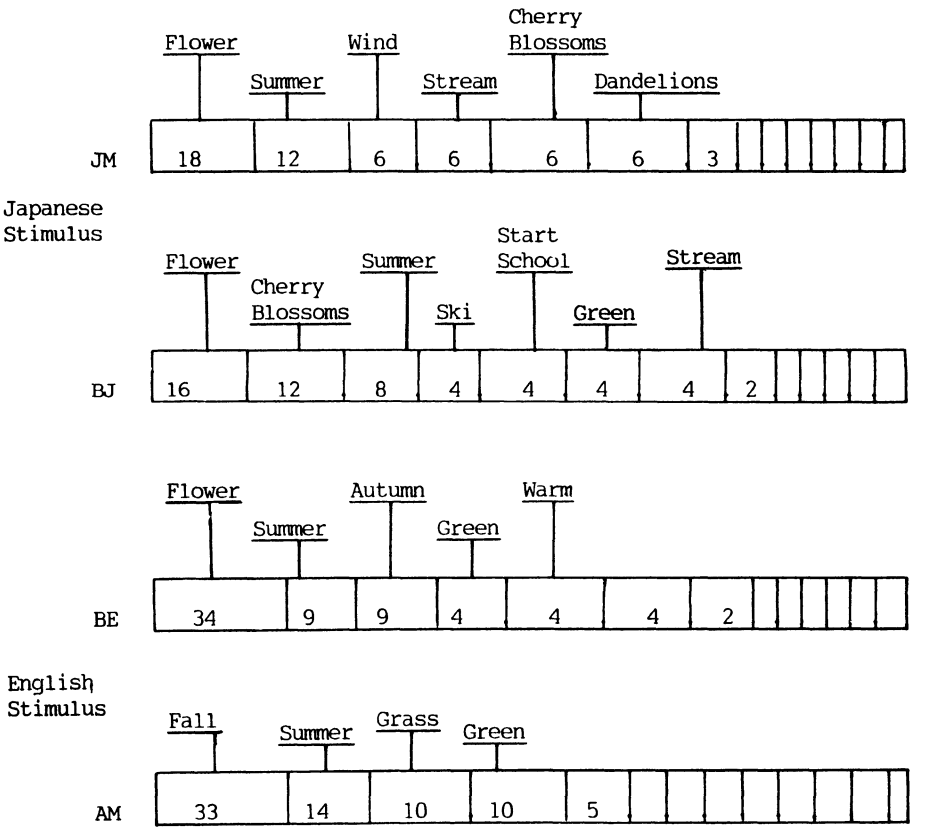


Table 1a. Haru - Spring

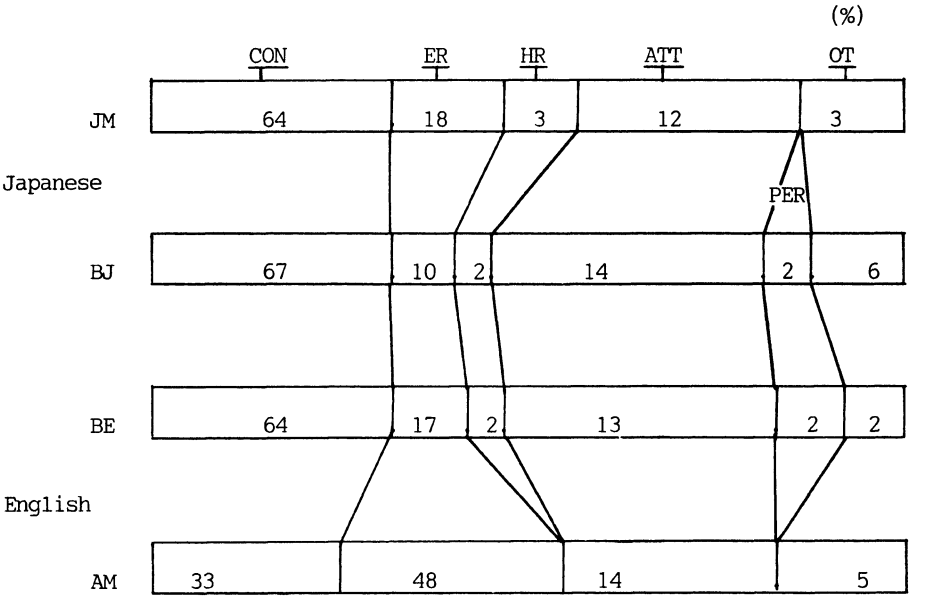


Table 1b. Haru - Spring

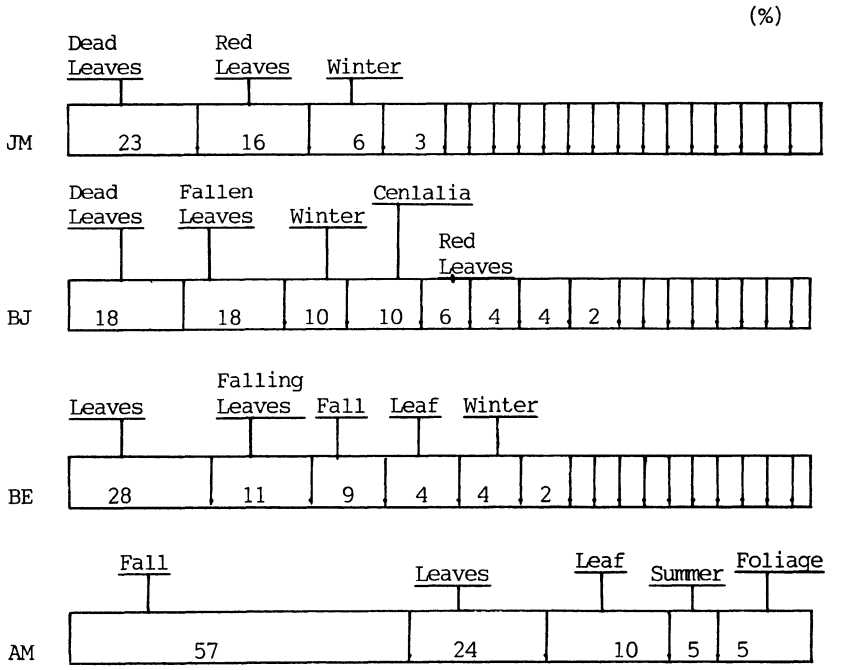


Table 2a. Aki - Autumn

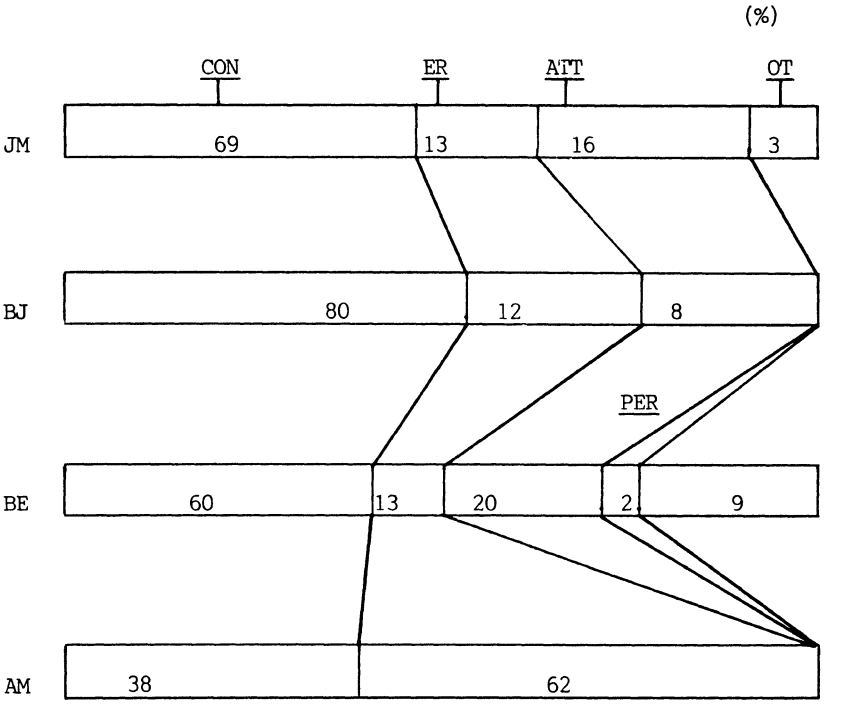


Table 2b. Aki - Autumn

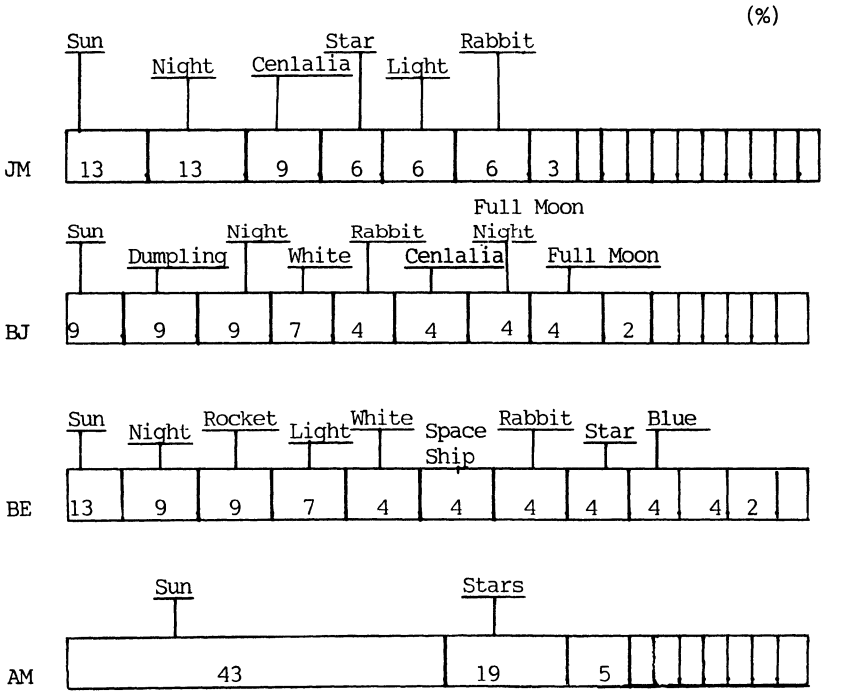


Table 3a. Tsuki - Moon

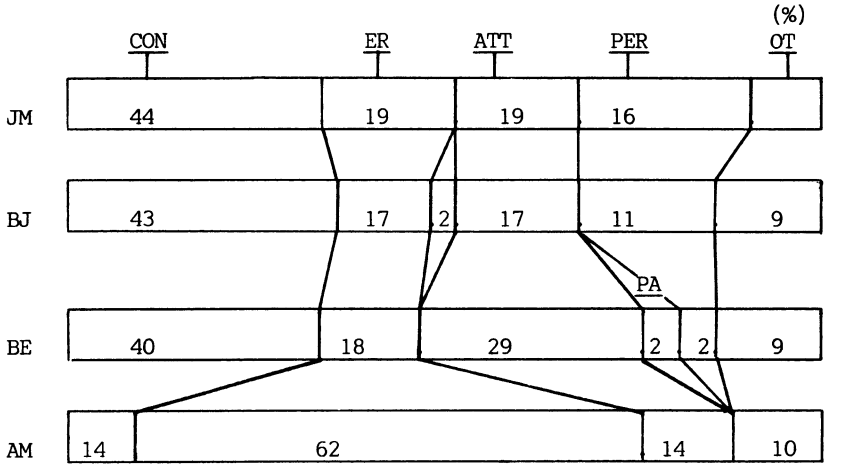


Table 3b. Tsuki - Moon

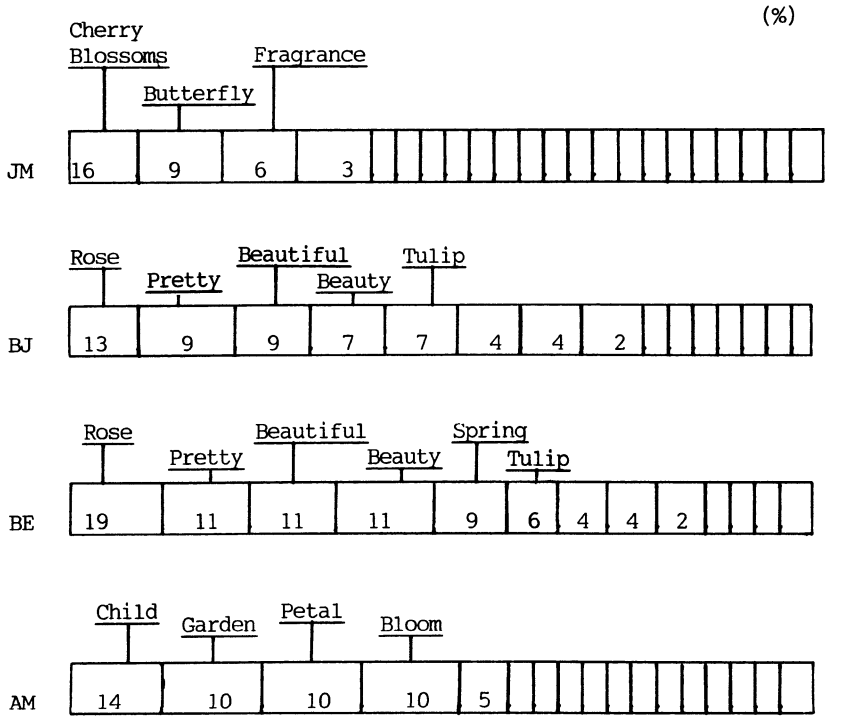


Table 4a. Hana - Flower

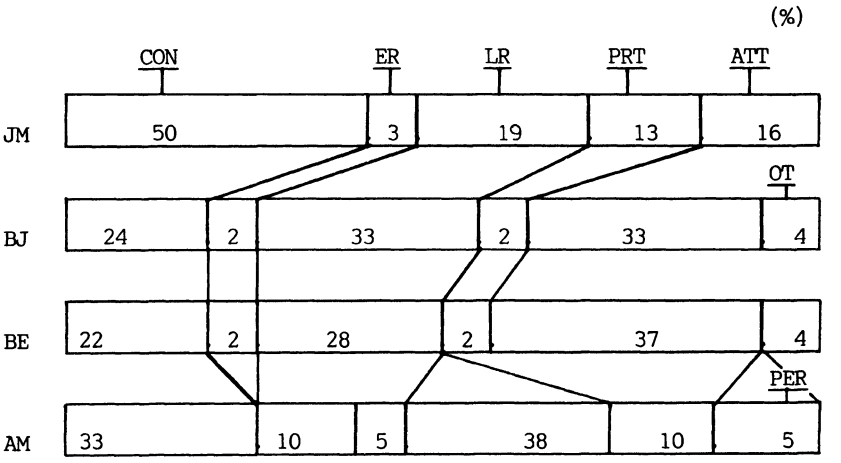


Table 4b. Hana - Flower

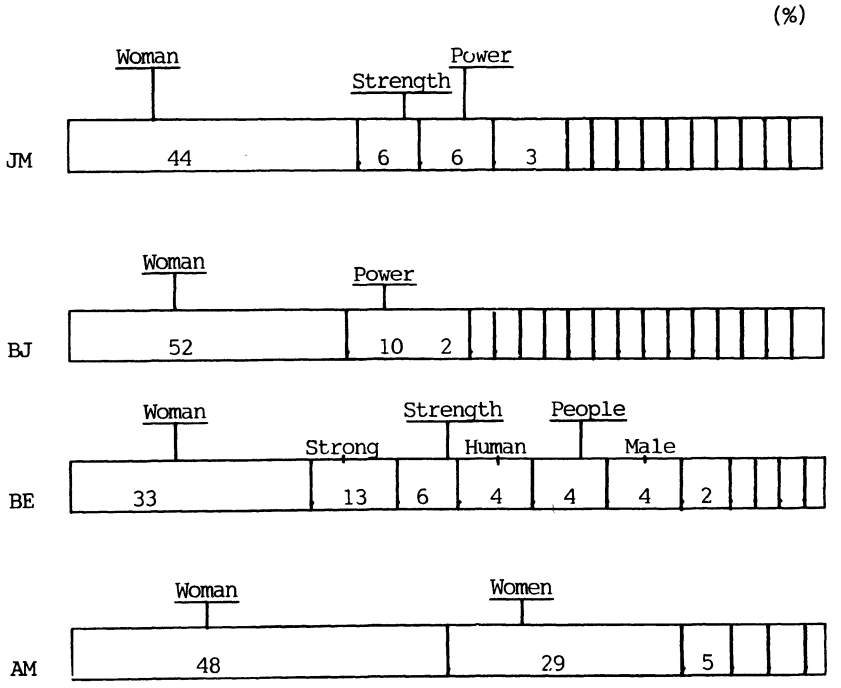


Table 5a. Otoko - Man

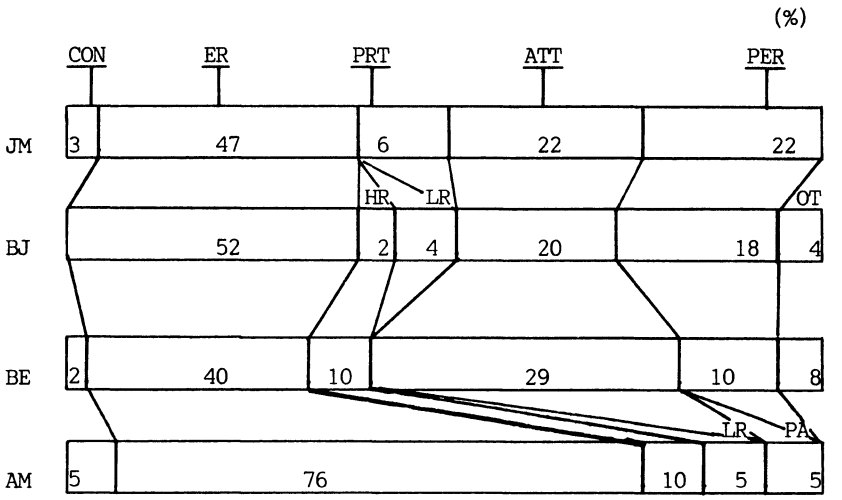


Table 5b. Otoko - Man

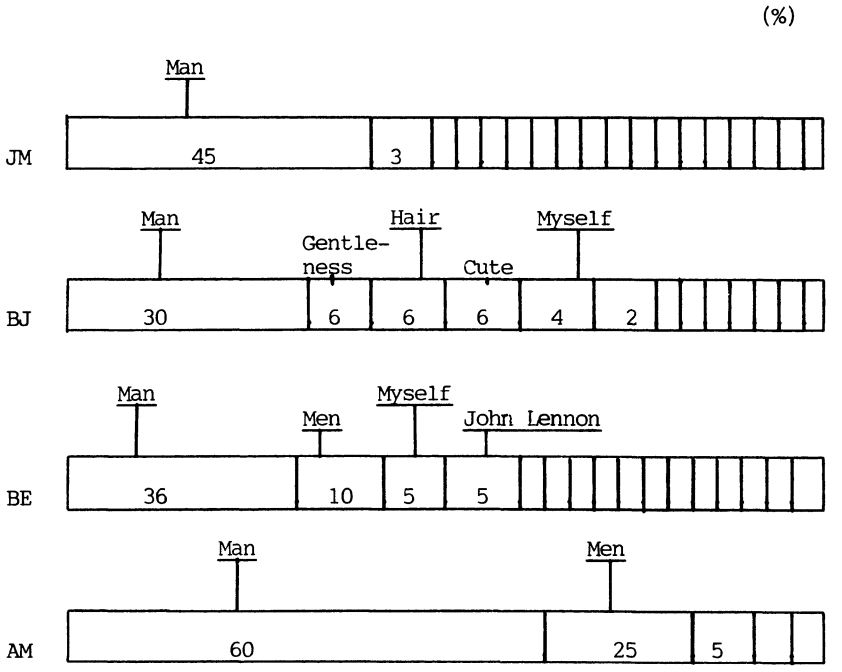


Table 6a. Onna - Woman

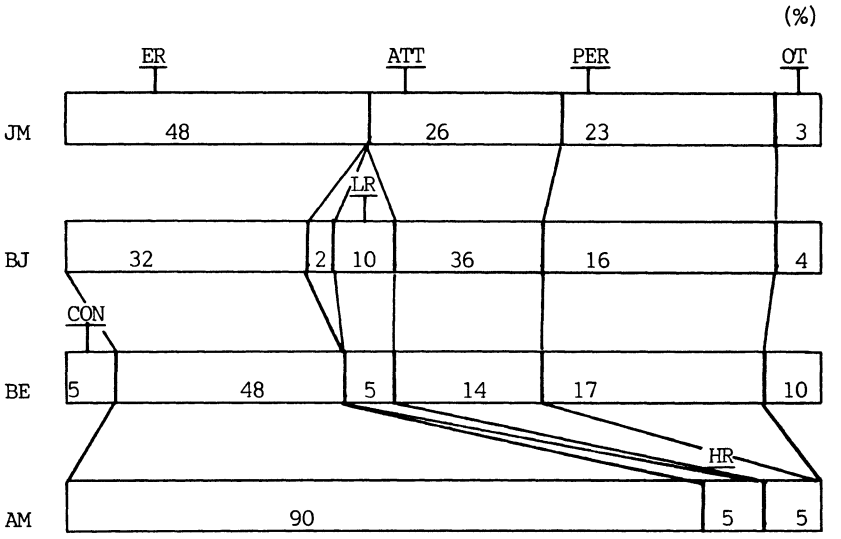


Table 6b. Onna - Woman

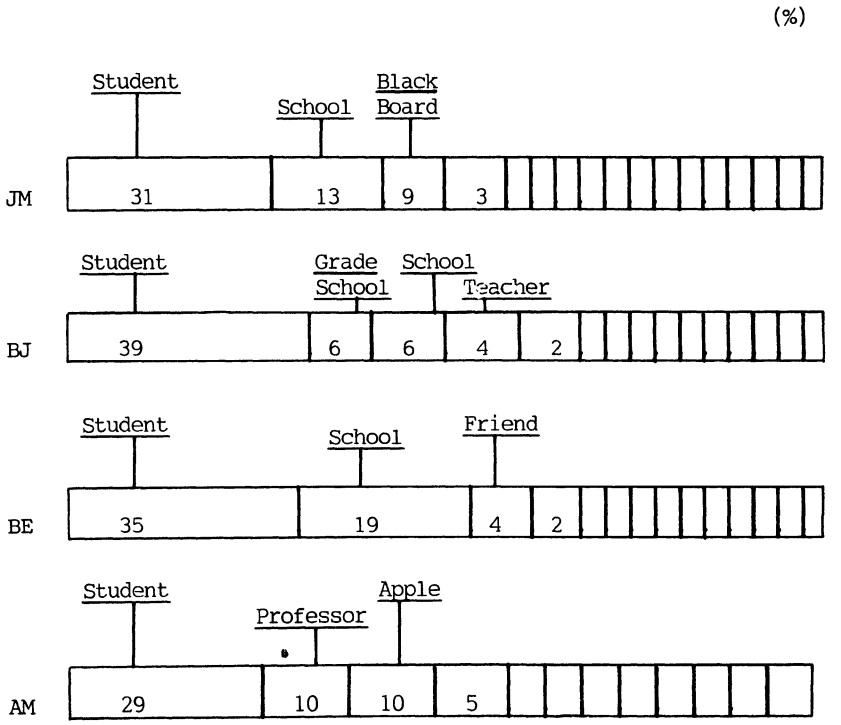


Table 7a. Sensei - Teacher

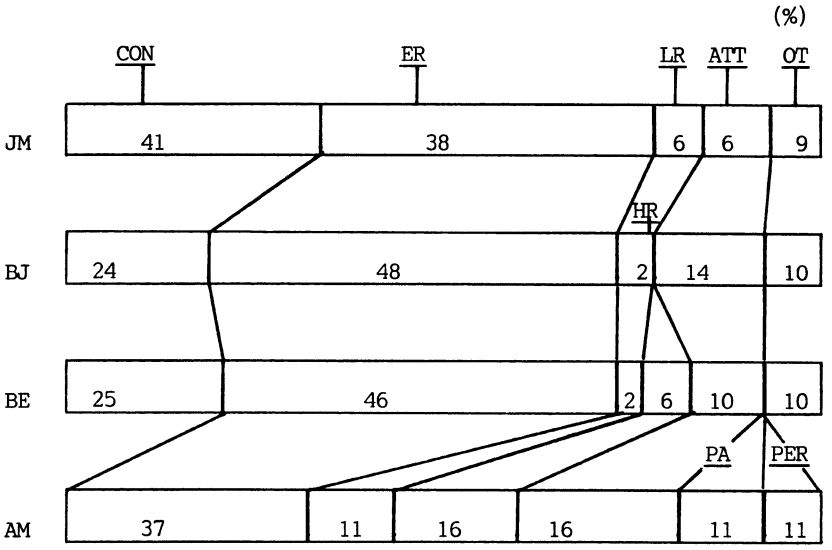


Table 7b. Sensei - Teacher

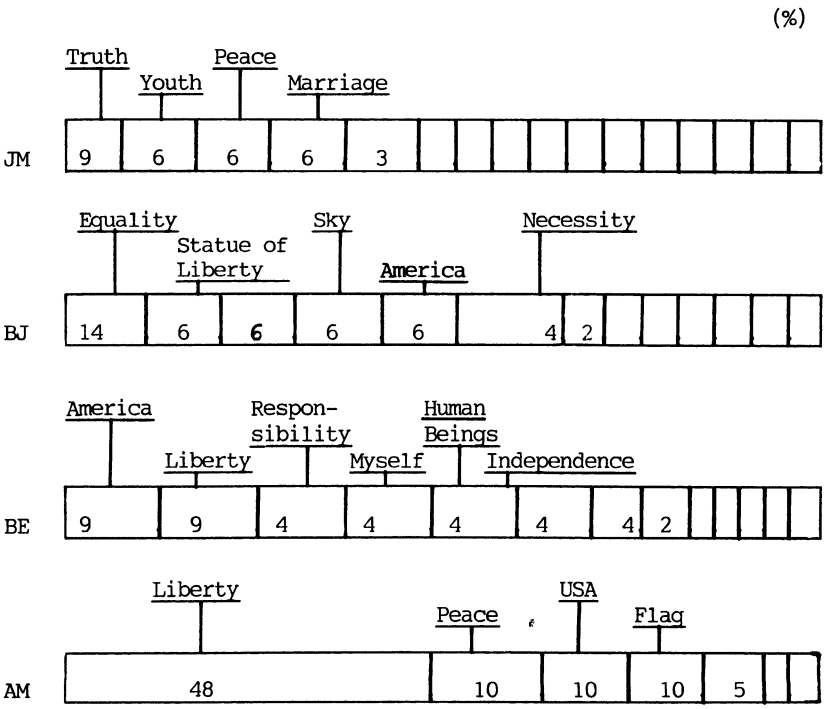


Table 8a. Jiyū - Freedom

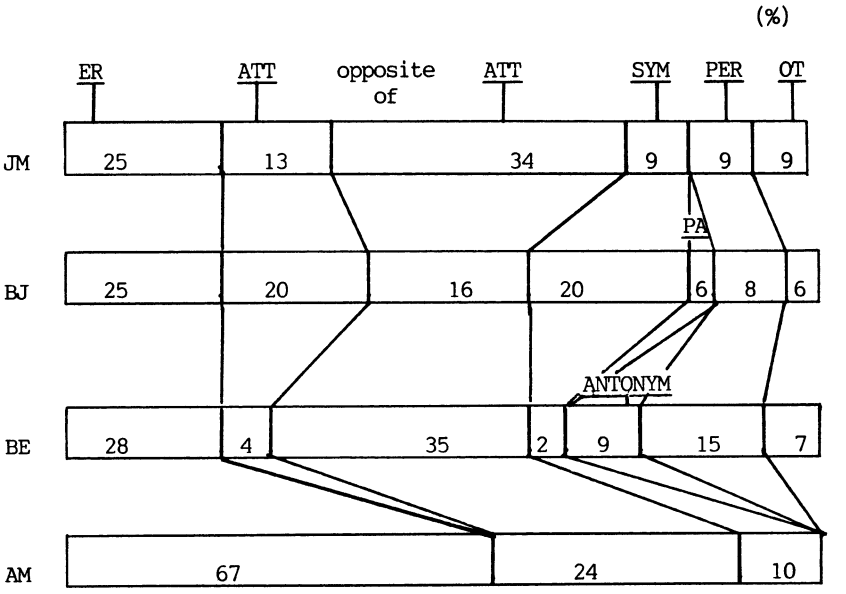


Table 3b. Jiyū - Freedom

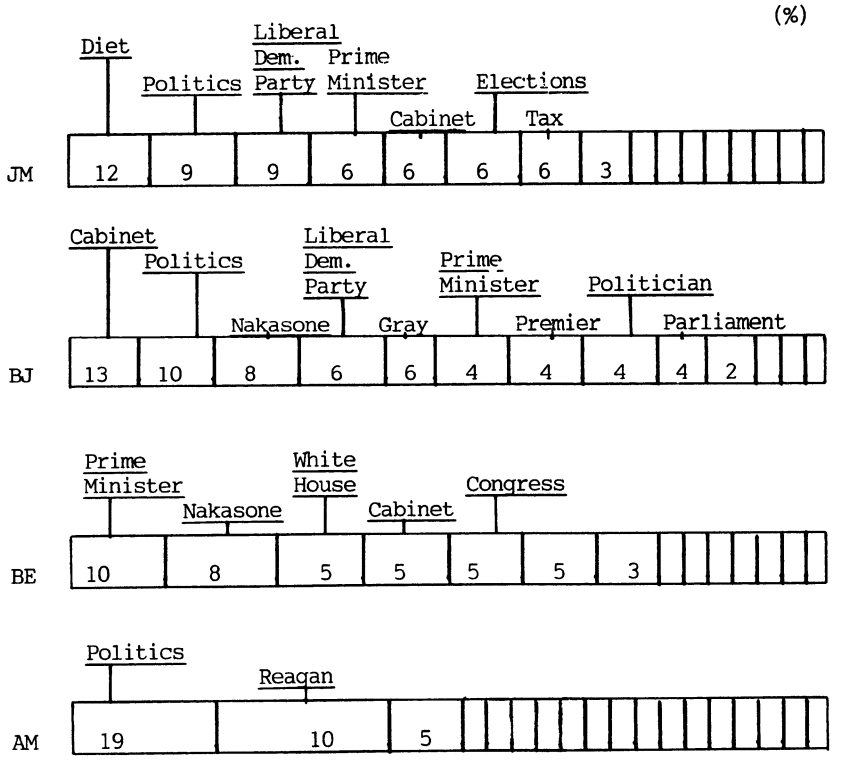


Table 9a. Seifu - Government

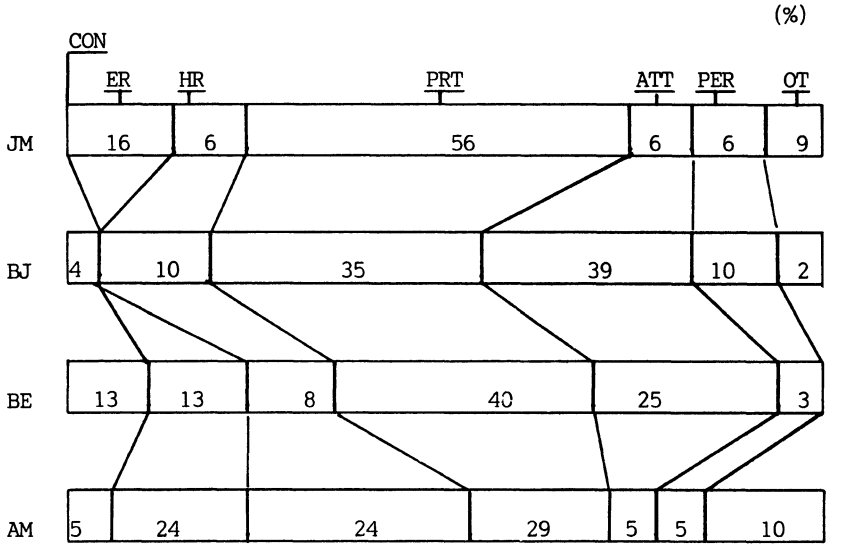


Table 9b. Seifu - Government

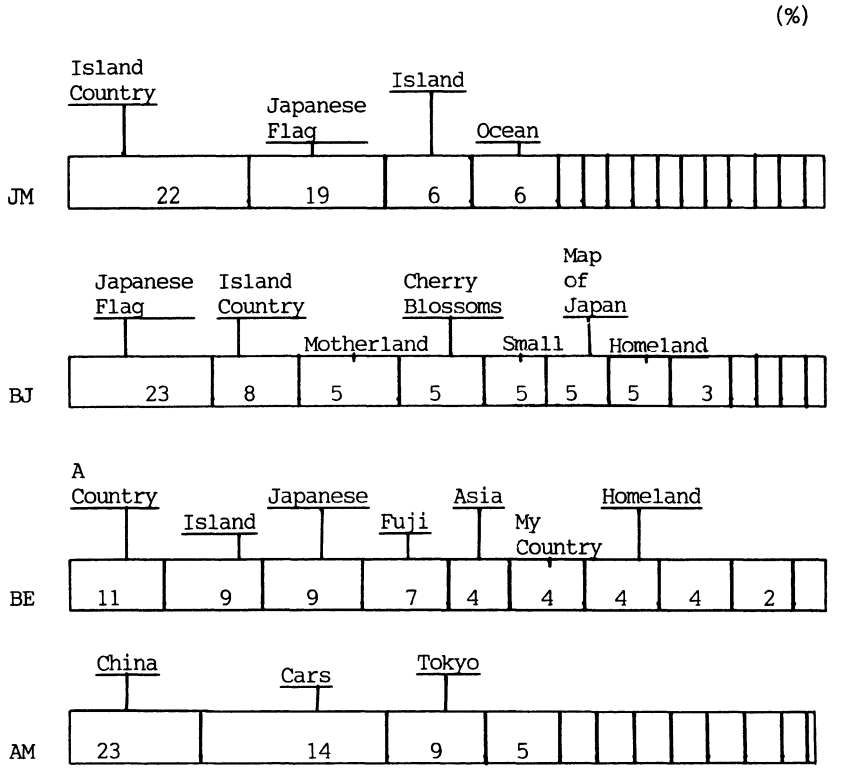


Table 10a. Nihon - Japan

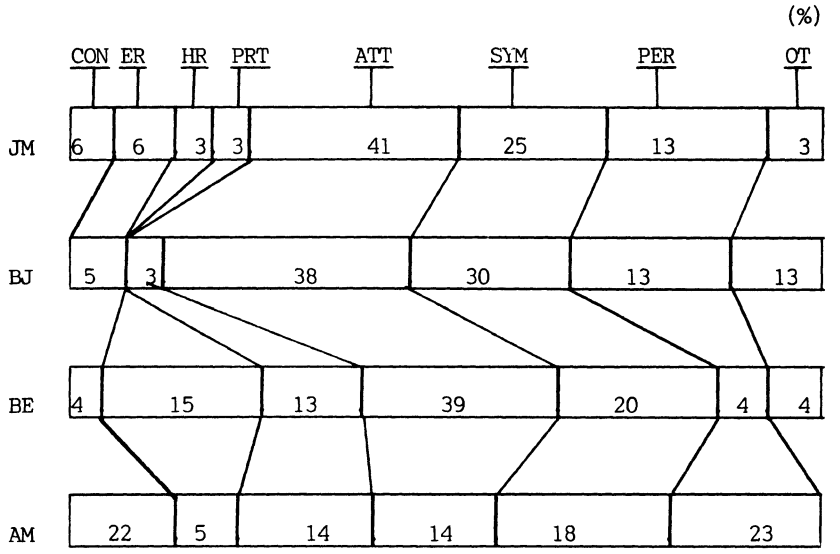


Table 10b. Nihon - Japan

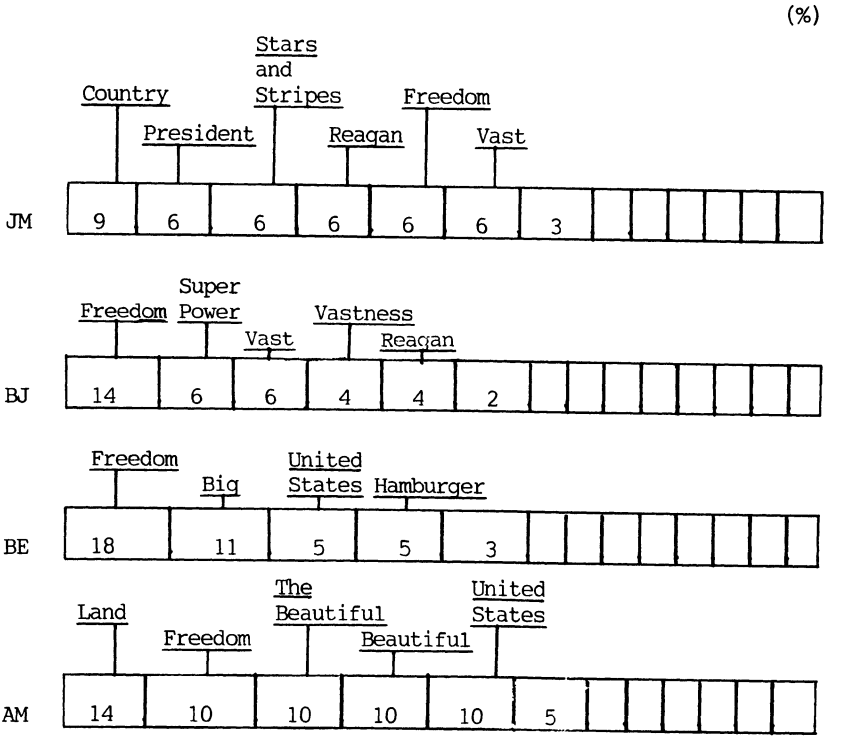


Table 11a. Amerika - America

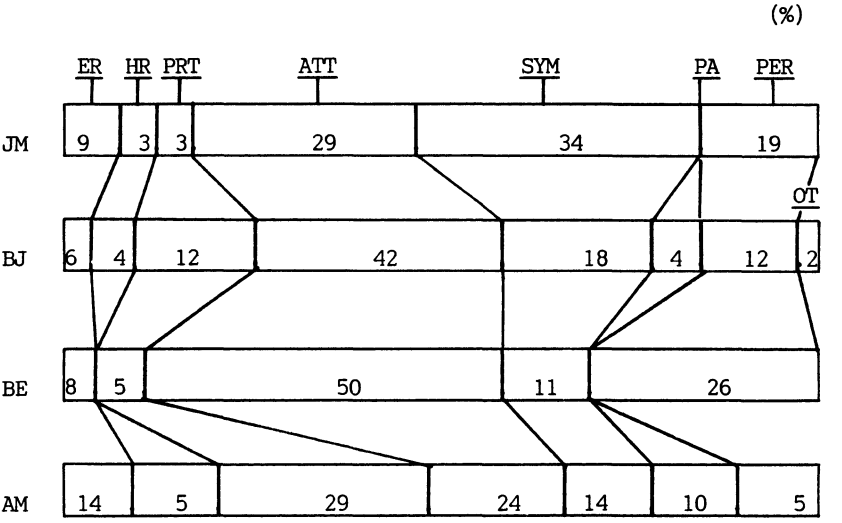


Table 11b. Amerika - America

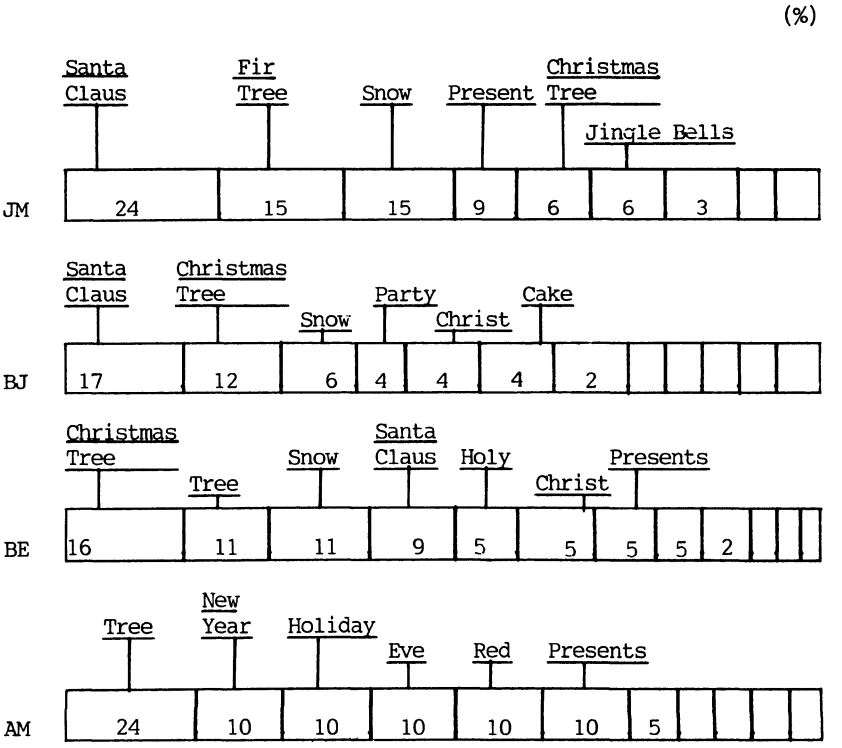


Table 12a. Kurisumasu - Christmas

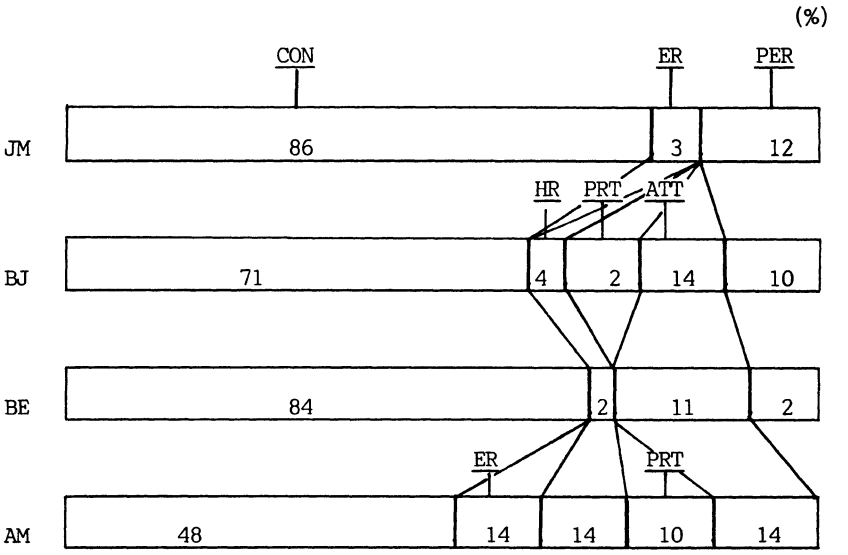


Table 12b. Kurisumasu - Christmas

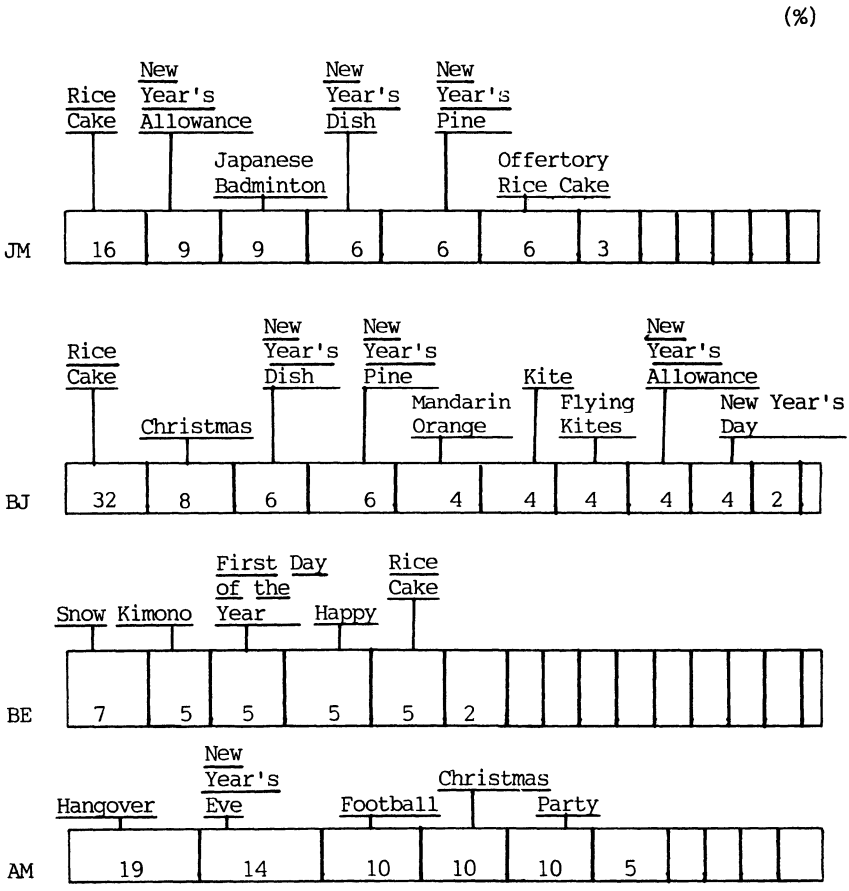


Table 13a. Shogatsu - New Year's Day

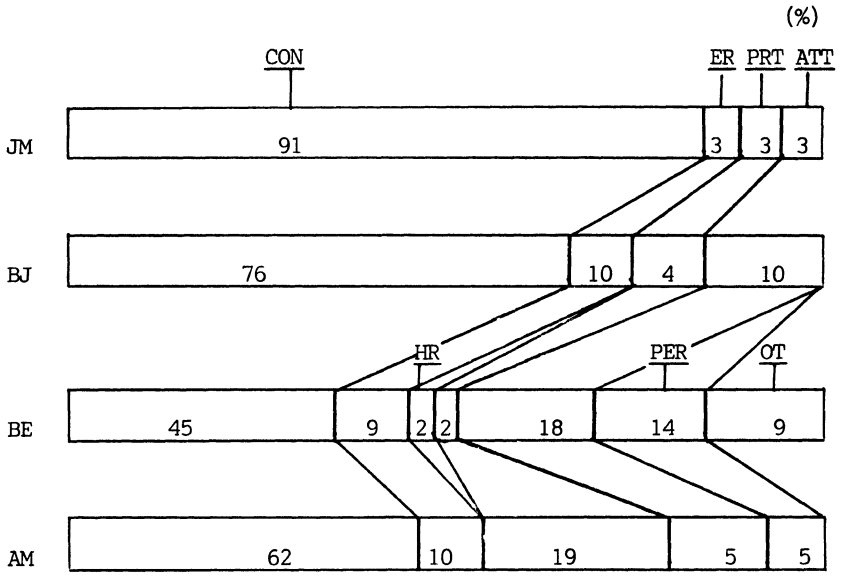


Table 13b. Shōgatsu - New Year's Day

Age of Leaving Japan	Length of Stay (Years)	Assimilation Type I	II	III	IV	V
Under 9	over 4	2			10	13 (FREQ.)
	under 4	1	4		9	
9 - 11	over 4	1		1	1	
	under 4	3	5	1	3	
11 - 14	over 4			4		
	under 4	2	4	2		
Over 14	over 4					
	under 4	1	2	1		
Subtotal	over 4	3		5	11	13
	under 4	7	15	4	12	
Total		10	15	9	23	13

(Minoura, Kodomono Ibunka Taiken (Tokyo: Shisakusha, 1984), Table 11-1, p.226)

Table 14. Japanese Children's Assimilation into American Culture

UNDERSTANDING INTERACTIVE COMPETENCE IN L1/L2 CONTRASTIVE CONTEXT:

A CASE OF BACK-CHANNEL BEHAVIOR IN JAPANESE AND ENGLISH

Senko K. Maynard

Rutgers University
New Brunswick, NJ

Recent studies in conversational interaction across speech communities have revealed that not only grammar proper, but also interactive strategies at a time of face-to-face encounter differ from language to language. Although in the past, conversation analysts have pointed out differences in some aspects of interactive competence, often such studies have been plagued by their ad-hoc nature of data, anecdotal accounts and casually defined contrastive context in which the analysis is made. In this study we attempt to understand one aspect of interactive competence, i.e., conversation management through back-channel strategies in a specific contrastive context of casual conversation in Japan and the United States. The methodology adopted here is what we may call "contrastive conversation analysis" as opposed to Hartman's (1980) "contrastive textology" in which primarily written text -- what he calls "parallel text"-- is analyzed. Note that although so-called contrastive analysis has been under attack over the last two decades -- primarily because of its failure to meet unrealizable expectation i.e., prediction of learners' errors in its own right, widely held among applied linguists especially in the United States-- the term contrastive analysis used herein is broader in its application and does not necessarily predict actual learners' errors. Based on this framework, this study investigates "back-channel" expressions, such as uh-huh's and brief comments received by the person who has the turn without relinquishing the turn as characterized first by Yngve (1971).

The data used for this study consists of twelve videotaped dyadic conversations, six American and six Japanese pairs, three male and three female in each country. The pairs selected in both countries were of the same sex and were all college students. In both countries the paired subjects identified each other as friends and they had had so-called casual conversation on numerous occasions. The filming was conducted in Tokyo and in New Jersey in May and February of 1985. In both locations an unattended video camera was used and the subjects were left alone in a room after they were instructed to talk about anything they liked as naturally as possible.

To minimize the influence of being videotaped reflected in our data, we categorically excluded the initial two minutes of film from our data and chose the three minute segments following the first two minutes as data relevant to this study.

To concentrate on back-channel clearly identified on the basis of the turn-taking context, in this study we focus on occurrences where an interlocutor who assumes primarily a listener's role sends brief messages during the other interlocutor's speaking turn. We refer to these back-channel occurrences as "turn-internal listener back channel."

Beyond verbalized uh-huh's and brief comments, various aspects of nonverbal behavior may potentially constitute back-channel expressions--such as direction and duration of gaze, use of self-adaptors, gesticulation and vertical and horizontal head movement. Because our data source fails to provide detailed eyegaze monitoring and extended occurrences of self-adaptors, and because head movement is observed frequently at the time when back-channels occur and it is speculated to play a crucial role, only clearly visible head movement is noted in this study. In observing head movement we have limited our focus on vertical head movement -- usually called "head nod" and horizontal head movement, the so-called "head shake."

Another utterable frequently observed in our data is laughter. As suggested by Schenkein (1972), laughter plays an important role in communication and in this study we identify the listener's laughter during the other interlocutor's turn to be a case of back-channel.

Having identified what specific phenomena we will be investigating in the present study, we now turn to a discussion of data. First let us observe data set (1), a 47-second segment of casual conversation taken from our data. (The capital H stands for head movement.)

(1) A: Dakedo (1)

are atsuryoku ga tsuyoi n da yo ne hora/ (2)

hoogakubu jan. (3)

(B:1 Aa sooka H hooka H uun H)

dakara (4)

mottai nai to ka iwarete sa H. (5)

B: A H mawari kara ne H. (1)

(A:1 H)

Oya kara sureba kodomo ga sureba iya LAUGH H (2)

(A:2 Soo H soo H soo soo)

A: Demo oya oya wa ne moo saikin soo mo (6)

(B:2 Soo)

iwa naku natta kedo H (7)

(B:3 H H)

tomodachi toka wa sa mottainai yoo to ka sa (8)

(B:4 Uun H)

Issho ni sa hora (9)

nihongo kyooshi no yoosei no kurasu no sa (10)

(B:5 Un)

tomodachi ttara nan da kedo issho ni uketeru hito ga

iru no. (11)

(B:6 Un un)

kyuu, juu nin no kurasu dakara (12)

(B:7 H Un H un)

moo shitashiku natte (13)

(B: 8H)

daitai toshiue no sa (14)

(B:9 H Uun un)

Ona no hito ga ooi wake H. (15)

(B:10 H Un)

moo ooeru yamete naru to ka ne. (16)

(B:11 H Un)

soo yuu hito ga sa (17)

B: 12 Un)

mottainai yo to ka yuu n da yo. (18)

B: Aa, sono W no hoogakubu made itte H. (3)
(A:3 Soo soo)

A: watashi nanka C datta no yo. (19)

B: Da dakara W W demo ne (4)

yappari naka de iru yo H ri (5)

(A: 4 Un)

maa ima no hyoogen de ichioo soo na n da kedo (6)

(A:5 Un)

soto kara mite mo H (7)

(A: 6H)

W no hoogakubu dattara doo ni demo naru maa

geneki tte yuu ka H. (8)

(A:7 H Soo H soo H soo)

English translation for data set (1):

A:1-3 But there's great pressure, 'cause I'm graduating from Law School.

(B:1 Oh I see, I see.)

A:4-5 So I'm told that it's not good enough for me.

B:1 You mean (you hear that) from people around you.

(A:1 Head movement)

B:2 From parents' view, if the child does

(A:2 Yeah, yeah, yeah.)

A:6 But nowadays parents don't

(B:2 I see.)

A:7 say those things.

(B:3 Head movement)

A:8 The way my friends look at it, they say things like, "It's not good enough for you."

(B:4 Yeah.)

A:9-10 You see we are all in the same Japanese language teacher training class.

(B:5 Yeah.)

A:11 and although we aren't exactly friends

(B:6 Uh-huh.)

A:12 There are nine or ten students in the class and

(B:7 Uh-huh.)

A:13 we have become friendly with each other.

(B:8 Head movement)

A:14 Most of them are older

(B:9 Yeah.)

A:15 women, many of them are.

(B:10 Uh-huh)

A:16 They wish to become teachers after leaving clerical positions at companies.

(B:11 Uh-huh.)

A:17 And those people

(B:12 UH-huh.)

A:18 say to me, "The job isn't good enough for you."

B:3 Oh, I see, they say (something like) you're graduating from W University's Law School.

(A:3 Yeah, yeah.)

A:19 "I graduated from C University you know," they say.

B:4-5 So although we are talking about W University, more than the way those of us on the inside can see it,

(A:4 Yeah.)

B:6 as you know what I mean by the expression I just used

(A:5 Yeah.)

B:7 when seen from the outside,

(A:6 Head movement)

B:8 the Law School of W University is still good and is first rate, I should say.

(A:7 Yeah, yeah, yeah.)

We observe in data set (1) 12 cases of B's back-channel during A's turn, and 7 cases of A's back-channel during B's turn. Some back-channel devices are strictly verbal as in the case of B's back-channel 2, soo; some are strictly nonverbal as in A's back-channel 1. Some listener back-channel behavior was a combination of verbal and nonverbal signs as in B's short utterance 1, Aa sooka hooka uun accompanied by three repetitious head movements. Our cursory observation reveals that back-channel strategies in Japanese involve various brief phrases with or without head movement. We also find that during the 47 second interval turn-internal listener back-channel occurred 19 times, at least once in every 2.47 seconds.

Now let us examine comparable data, 47 second segments taken from our American subjects, data set (2).

(2) A: I ordered some escargot (1)

and got me a coke (2)

I was like (3)

B: I have never been to K. Miller. (1)

A: I don't know just like (4)

strikes me as being very pseudointellectual (5)

Don and I were walking past ( ) going to that

little shop (6)

past it's open only three days or

something. (7)

(B:1 Um Hum)

you know the one I bought my uh (8)

dice bag. (9)

B: Yeah I think I know what you mean. (2)

(A:1 Yeah)

A: and we were going there and this guy came out

of K. Miller because he notices us looking at

the menu and he goes (10)

Hey, Babe, want a drink? Come on inside I'll

pay for you. (11)

(B:2 LAUGH)

And we were like "Oh go away." (12)

B: Weird (3)

(A:2 Yeah)

No I heard the food's actually good though. (4)

A: All I know is Polly offered me a slimy little

escargot and I said thank you but no. LAUGH (13)

B: Oh I like escargot. (5)

A: I don't (14)

I I just keep on thinking slime (15)

sledge (16)

sea bottoms, you know (17)

We observe only four cases of back-channel in our data set (2), two by speaker B and two by speaker A. We see three utterances and one case of LAUGH, none of which is accompanied by head movement. Our observation of Japanese and English data, although very preliminary at this point, reveals that although back-channel phenomena exist in both languages, they utilize different devices and operate in a different manner and frequency.

After examining 36 minutes--three minutes each from 12 pairs--of

Table 1
Frequency of turn-internal listener back-channel
among Japanese and American pairs

Japanese pairs:		American pairs:	
1	45	1	22
2	37	2	8
3	35	3	13
4	42	4	15
5	39	5	6
6	29	6	9
Total	227 (mean=37.83)	Total	73 (mean=12.17

conversational data, we observed a significant difference in frequencies of back-channels. The overall frequency count for the turn-internal listener back-channel was 227 among Japanese speakers. In English data we find listener back-channel behavior significantly less frequently, totalling only 73 times. In all Japanese pairs we observe higher frequency of back-channel than in any of the American pairs. Table 1 lists the observed frequency of back-channel occurrences in each of the Japanese and American pairs with the mean for each group.

The most frequently occurring types of back-channel strategy, brief utterables, totaled 58% for Japanese and 48% for English. Head movement accompanied these brief utterables 63% of the time in Japanese but merely 14% among American subjects. In the second most frequent category, vertical head movement (which does not accompany verbal cues), Japanese pairs produced it 76 times while American subjects did so 15 times.

In English data we observe that brief phrases and ellipted sentences such as I don't remember occur 10 times (14%) whereas in Japanese such phrases occur 13 times and constitute only 5%. The laugh category was observed 30 times and 13 times among Japanese and American pairs respectively. Horizontal head movement occurred once in Japanese data, while the figure in American data was 4.

Having discussed the frequency of back-channel, let us now focus on function. Why do listeners feel a need to send back-channel signals in the first place?

According to Schegloff (1982), back-channel is characterized as a "continuer" a sign through which the listener encourages the speaker to continue to talk. Schegloff reaches this interpretation on the basis of the fact that the listener passes up opportunities for other-initiated repair during moments in which the listener could potentially take turns.

Note that this characterization by Schegloff was derived on the basis of the notion of sequencing in conversational interaction. In the

tradition of ethnomethodological conversation analysis, the assumption taken is that, as Atkinson and Heritage (1984:7) put it, "turn within sequence character provides a central resource for both the participants and the overhearing analyst to make sense of the talk." In other words, it is the sequence of interaction and not the observation obtained otherwise that provides clues for interpreting the function of back-channel.

The notion of sequencing and the procedure for investigating conversational turn in sequence, however, although useful as analytic devices, sometimes suffer from over-simplification which may obscure important features of interactional complexity of situated talk. Note that this characterization of sequencing fails to capture back-channel behavior that co-occurs with the speaker's verbalization when there is no pause. Similarly, back-channel behavior may sometimes occur concurrently with the last syllable of the speaker's utterance. These cases of back-channel activities are used while the speaker continues to talk without clearly hinting that the speaker is going to pause.

This implies that back-channel in Japanese, at least in the data examined, is used in an environment where other-initiated repair is not expected. Thus we can not characterize back-channel simply as a "continuer." Passing up an opportunity for repair becomes irrelevant since such opportunity never existed in the first place. In other words, although back-channel behavior may function in part as a "continuer" and may be characterized as an expression of passing up an opportunity to initiate repair as suggested by Schegloff, this characterization does not account for all cases.

One solution to this problem is incorporating various functions on the basis of different aspects of communication. We propose that functions of the back-channel behavior in Japanese and English may be specified in terms of the following five categories although they are not meant to be exhaustive nor mutually exclusive.

1) continuer

2) display of understanding of content

3) support and empathy toward the speaker

4) agreement

5) strong emotional response

These functional characteristics are derived from various aspects of conversation. First, "continuer" is justified on the basis of a turn-taking system and specifically on the basis of foresaking the opportunity for repair as proposed by Schegloff. Displaying an understanding of content is identified when confirmation of the listener's understanding is felt necessary. One such example is observed in B's back-channel (1) in data set (2). Naturally sending back-channel does not guarantee real understanding of the content. It merely expresses the listener's claim that s/he understands. Support and empathy for the speaker is observed in those instances where the speaker makes an evaluative statement to which support or empathy is felt necessary by the listener. Such is the case when A gives back-channel (1) in data set (2). Agreement is identified when the speaker's turn performs a speech act of questioning or question-like statements. Back-channel here performs a specific speech act of agreeing, and this is done without the speaker relinquishing the turn. For example A's back-channel (1) in data set (1) may be interpreted

as mild agreement. Strong emotional response is identified when the listener sends exclamatory phrases or laughs which indicate more than simple continuer, understanding or support. These functions do overlap and it is best understood that we identify each turn-internal back-channel as performing more or less one of the functions proposed.

Let us now focus our attention on another issue relating to back-channel expression, the context in which back channel behavior occurs. In Japanese, first, we observe that in the majority of cases turn-internal listener back-channel behavior occurs during the speaker pauses immediately following the speaker's utterance.

Secondly we observe that these pauses are marked by certain linguistic devices that the speaker frequently selects. One such device is the so-called final particles (shuujoshi and kantoojoshi). In data set (1), particle ne occurs three times while sa occurs five times and yo twice, and no and ka each occurs once at the point of clausal unit. Final particles in Japanese (or what Martin 1975 calls "sentence extension") are normally used to impact some additional hint of the speaker's attitude toward what s/he is saying. Ne, which is often translated into English tag questions and what Uyeno (1971) calls a "particle of rapport," solicits listener response--either in the form of the back-channel (if the speaker continues the turn) or in the form of a change of turn.

Particle endings in general then seem to signal, together with the pause itself, the moment where some sort of feed back--which is most likely to be in the form of back-channel--may be relevantly performed in the sequence of the conversation-in-progress. Particle endings marked approximately 50% of all back-channel occurrence.

Thirdly, we find that even when no particles are attached, pauses occur at the major clause junctures--such as at the juncture of coordinate and subordinate clauses. Note first that subordinate clauses which are marked by conjunctions such as ba in B's (2) in data set (1) also function to mark the ending of an utterance. In such cases the matrix clause may not appear at all. As suggested by Ooishi (1970), although these subordinate conjunctions serve conjunctive functions, they appear in casual conversation in ellipted sentences, where the main clause is not expected. Such ending leaves some space for the listener to fill in his/her own interpretation of the main clause and thereby provides interactional options for both participants. In understanding the context in Japanese conversation which constitutes relevant cues for back-channel, we must incorporate our somewhat extended notion of the grammatical unit where conjunctive particles may indeed mark its end. Among 6 pairs of the subjects examined, out of 227 locations where back-channels occurred, 71% occurred at major grammatical junctures, some accompanied by sentence final particles and/or head movements.

Another point of interest, gerundive verb endings were found to mark utterance boundaries. For example, natte in A's (13), and itte in B's (3) in data set (1) mark utterance boundaries. The gerundive form is used as the conjunction 'and' gives the impression that the talk is not completely over and therefore provides the listener an opportunity to respond with a "continuer."

Data set (1) also reveals an interesting nonverbal behavior on the part of the speaker at the point where turn-internal listener back-channel is observed, namely, the head movement that cooccurs at or near the final syllable of the speaker's utterance. There are seven such cases where back-channel devices follow speaker's head movement--(in B's (2) A's (13), A's (15), B's (3), B's (5), B's (7) and B's (8)). These

head movements were observed quite frequently across all data we examined. Out of a total 227 back channels, 28% occurred in the context of the speaker's head movement.

In summary, regarding the context for back channel, recurring patterns emerge; 1) usually the listener back channel occurs during the pause between clausal units of utterance within the speaker's turn, 2) these clausal units are marked not only by standard sentence ending forms such as verbs but also by gerundive forms of the verbs and conjunctive particles, 3) the back-channel often appears at these clausal boundaries marked by final particles such as ne, sa and yo, and finally, 4) the sentence final syllable which precedes listener back channel may be accompanied by head movement.

In terms of the types of context and their association to the back-channel strategy in American English, we focused on the devices similar to what we examined in Japanese, namely, grammatical completion, phrase ending markers such as you know (what Bernstein calls "sympathetic circularity sequences), and the head movement.

Among 12 American subjects we observe 97% of back-channel at the point of grammatical completion. Sympathetic circularity sequences mark the context for back-channel in only 5 cases of the 73 back-channels (7%), while head movement occurs only in 6% of the total locations of back-channel. In English, then, the grammatical completion point is the single most powerful discourse context for back-channel; other criteria appear to mark only marginally the points of relevance for back-channel.

We can conclude then, that although in both Japanese and English, grammatical completion is a relevant discourse context, in Japanese sentence particles and head movements also mark to a reasonably significant degree the context for back-channel. Note, however, that this study does not address the predictability of back-channel occurrences given the discourse contexts identified above as providing "cues." We have only observed actual occurrences of back-channel expressions and the characteristics of their discourse contexts.

Before we proceed, one point of caution should be noted. We have only examined devices similar to what are observed in Japanese context for contrastive purposes. Naturally both in Japanese and English, there are other potential signs that are candidates for relevant context. Therefore, we must be warned that our conclusions are based on these limited select devices under discussion here. However, in general it is reasonable to conclude (in regard to three devices examined in this study) that English does not provide, above and beyond grammatical completion, overt devices such as the types available in Japanese to encourage listener back-channel responses as overtly or as frequently as in Japanese.

What we observed in this study implies that "linguistic relativity" exists not only in lexicon and syntax but also in the ways people manage conversation interactionally. Although it is unwise to assume, without reasonable supporting evidence, that there exist correlations or general correspondences between linguistic norms and its contextual cultural norms, we can make the following interpretations.

The continuous flow of back-channels sent by Japanese listeners, and the speaker's ready acceptance of such extraordinarily frequent feedback in comparison to American English, suggest that Japanese interactants possess a strong indication for mutual monitoring and cooperation. Whereas points similar to this have been made by Clancy (1982) and Mizutani (1983), among others, such characterization has been made frequently in an imagined

or casually defined contrastive context. The result of this study designed in the contrastive context of Japanese and American English provides some evidence to support this long-claimed preoccupation of the Japanese people, i.e., conflict-avoiding, cooperative style of social interaction.

The differences in interactive competence in L1 and L2 as observed in this study offer suggestions toward pedagogical applications. For example, language learners may be encouraged to focus on interactive aspects of conversation--rather than to merely pay close attention to the language of conversation as we usually do. Instructional videos and/or self-produced videos of conversational interaction may be used in order to provide language learners an opportunity to study so-called "conversation management." Students may be directed to make note of: (1) how to start and end conversations, (2) how the strategies for the turn-taking system are realized, (3) how to send back-channel expressions, (4) how to introduce a new topic and how to develop it interactionally, and above all (5) they may be encouraged to understand the interactional style that is socioculturally defined in the target speech community.

As a final note we should mention some drawbacks of this study. First we artificially excluded the participant's behavior during the inter-turn pause, which eventually has to be incorporated before we reach the final conclusion. Secondly, our data is limited not only in numbers but also in terms of genre. Whether or not what we discovered in this study of casual conversational discourse is applicable to other genres awaits further investigation.

And thirdly, when we undertake a contrastive analysis as designed in this study we must face the critical issue of "equivalence." In order to conduct a contrastive discourse analysis within the context of contrastive analysis such as proposed by Whitman (1970), one must first obtain "equivalent" forms in languages to be contrasted. What constitutes "equivalence" in each society contrasted is controversial and remains unresolved. Our attempt to obtain equivalence in this study is made in terms of the type of genre and the devices investigated. Thus we must remain cautious as we evaluate the results of this study which is based on a narrow and limited perspective.

*Portions of this paper have appeared in an article by the same author titled "On back-channel behavior in Japanese and English" in Linguistics, Vol. 24, No. 6, pp 1079-1108 (1986).

REFERENCES

Atkinson, J. Maxwell and Heritage, J. (eds.) (1984). Structure of social action. Cambridge: Cambridge University Press.

Bernstein, B. (1962). Social class, linguistic codes and grammatical elements. Language and Speech, 5, 221-240.

Brunner, L.J. (1979). Smiles can be back-channels. Journal of Personality and Social Psychology. 37,5, 728-734.

Clancy, P. (1982). Written and spoken style in Japanese narratives. In D. Tannen (ed.) Spoken and written Language, 55-76. Norwood, N.J.: Ablex.

Duncan, Starkey and Donald Fiske. (1977). Face-to-face interaction: research, methods, and theory. Hillsdale, N.J.: Lawrence Erlbaum.

_____. (1985). Interaction structure and strategy. Cambridge: Cambridge University Press.

Erickson, F. (1979). Talking down: some cultural sources of miscommunication in interracial interviews. In A. Wolfgang (ed.) Nonverbal behavior: applications and cultural implications, 99-126. New York: Academic Press.

Goodwin, Charles. (1981). Conversational organization: interaction between speakers and hearers. New York: Academic Press.
Hartmann, R.R.K. (1980). Contrastive textology. Heidelberg: Julius Groos Verlag.
Hinds, J. (1978). Conversational structure: An investigation based on Japanese interview discourse. In J. Hinds and I. Howard (eds.), Problems in Japanese syntax and semantics, 79-121. Tokyo:Kaitakusha.
_____. (1982). Japanese conversational structures. Lingua, 57, 301-326.
Labov, William. (1972). Sociolinguistic patterns. Philadelphia: University of Pennsylvania Press.
Martin, Samual, (1975). A reference grammar of Japanese. New Haven: Yale University Press.
Mizutani, N. (1983). Aizuchi to ootoo., In O. Mizutani (ed.) Hanashi kotoba no hyoogen, 37-44. Tokyo: Chikuma Shoboo.
Ooishi, H. (1970) Hanashi kotoba to wa nani ka. Hanashi kotoba shiriizu 12 by Bunka-cho, 36-51. Tokyo Ookurashoo Insatsukyoku.
Sacks, H. E. Schegloff and G. Jefferson. (1974). A simplest systematics for the organization of turn-taking for conversation. Language 50, 696-735.
Schegloff, E. (1968). Sequencing in conversational openings. American Anthropologist 70, 1075-1095.
_____. (1982). Discourse as an interactional achievement: some use of "UH-HUH" and other things that come between sentences. In D. Tannen (ed.) Georgetown University Round Table on Languages and Linguistics, Analyzing discourse: Text and Talk, 71-93. Washington, D.C.:Georgetown University Press.
Schenkein, J.N. (1972). Towards the analysis of natural conversation and the sense of heheh. Semiotica 6, 344-377.
Whitman, R. (1970). Contrastive analysis: Problems and procedures. Language Learning 20, 191-197.
Uyeno, T. (1971). A study of Japanese modality--A performative analysis of sentence particles. University of Michigan dissertation.
Yngve, V.H. (1970). On getting a word in edgewise. Chicago Linguistics Society, 6th, 567-578.

USING COGNITIVE GRAMMAR, NATURAL PHONOLOGY, AND ITEM RESPONSE THEORY TO EXPLAIN ESL SUBJECTS' CONTROL OF THE [D] and [Z] MORPHEMES

Kyle Perkins and Sheila R. Brutten

Southern Illinois University

Carbondale, Illinois

INTRODUCTION

Our purpose in this paper is to apply selected notions from cognitive grammar, natural phonology, and item response theory in order to give better insight on ESL subjects' processing of the D and Z morphemes in English.

According to Nathan (1986) cognitive grammar (formerly sometimes known as space grammar) is a theory which has the intention of integrating human cognition, perception, and language into a more coherent view of the relationship between language and mind. At the moment Ronald Langacker and George Lakoff are the leading proponents of cognitive grammar. The aim of this approach, according to Nathan, is to define linguistic functions in cognitive terms. Nathan further explains the intention of these two linguistic approaches: "language structure in general is not significantly different from other aspects of human cognition, and ...constraints on the nature and functioning of cognition are reflected directly in the nature and functioning of human language" (p.3).

The following quote from Langacker (1982) details the units of description in cognitive grammar:

> The grammar of a language, in space grammar terms, is simply a structured inventory of conventional linguistic units. For a given speaker, this inventory defines the boundaries of established linguistic convention as he knows it. Schematic units coexist in the grammar with content units that elaborate them; their relation is dynamic and interactive. On the one hand, a schema commonly derives from an array of previously established content units; extracting the schema amounts to perceiving the similarities which unite these units and abstracting away from their points of divergence. It should be noted that schemata are the only abstract descriptive units posited in space grammar. They are the functional equivalent of rules, and embody generalizations; but they are constrained in that they must always and only bear schematic relations to content structures (p. 25).

The thesis of Nathan's paper is that cognitive grammar nicely accommodates certain concepts in natural phonology, namely the notion of mental representation as physical space or embodiment and the notion of categorization by prototype. According to Nathan, natural phonology focuses on the mental nature of linguistic structures and any phonological processes to be included in a theory of natural phonology are intended to be real and to represent actual mental events. Phonemes are viewed as mental sound images which speakers perceive their language being composed of and as articulatory targets which speakers believe they are producing. Phonemes, then, are real sounds and embodied representations in mental space which combine the speaker's knowledge of the acoustic signal and the articulatory muscular movements which are necessary to produce the sound waves. A particular phoneme is composed of a class of sounds called allophones. All these allophones are classed as the same in terms of their commonality or their degree of similarity as prototypical members of a phoneme. The notion of the prototypical effects of allophones stems from the research of Eleanor Rosch (1973, 1977, 1978). Prototypical effects for the organization of allophonic and phonemic units in language have been demonstrated by Nathan, and Rosch has demonstrated their usefulness in the study of general cognition.

Using Nathan's work as a background, we propose in this paper the hypothesis that the morphemes D and Z are stored as mental embodiments of sounds, as articulatory/perceptual images. We further hypothesize that these morphemes are categorized as prototypes with their allomorphs conceptualized as a population of objects which can be considered in terms of their deviations from the prototype. The phonologically conditioned choice of allommorphs is accounted for by lenition, one class of phonological processes proposed by natural phonology. Lenition includes assimilation, deletion, and other related allomorphic processes that ease the transitions between contiguous sounds, that define permissible deletions of phones, and that make language pronounceable.

Our basis for these hypotheses is forgrounded in cognitive grammar and natural phonology. Natural phonology holds phonemes (and morphemes like D and Z) are mental images and embodied in mental space. Cognitive theorists believe that phonology is organized along the same lines as language in general: language is part of the human mind and is organized through central concepts such as physical imagery and prototype effects in categorization.

For the data analysis portion of the paper we have used item response theory (IRT). Hambleton and Swaminathan (1985) give a succinct definition of IRT.

> Item response theory postulates that (a) examinee test performance can be predicted (or explained) by a set of factors called traits, latent traits or abilities, and (b) the relationship between examinee item performance and the set of traits assumed to be influencing item performance can be described by a monotonically increasing function called an item characteristic function. This function specifies that examinees with scores on the traits have higher expected probabilities for answering the item correctly than examinees with lower scores on the traits. In practice, it is common for users of item response theory to assume that there is one dominant factor or ability which explains performance. In the one-trait or one-dimensional model, the item characteristic function is called an item characteristic curve (ICC) and it provides the

> probability of examinees answering an item correctly for examinees at different points on the ability scale.
>
> The goal of item response theory is to provide both invariant item statistics and ability estimates. These features will be obtained when there is a reasonable fit between the chosen model and the data set. Through the estimation process, items and persons are placed on the ability scale in such a way that there is as close a relationship as possible between the expected examinee probability parameters and the actual probabilities of performance for examinees positioned at each ability level (pp. 13-14).

Of particular interest in this paper is the term latent trait or latent space. Latent trait or latent space can be defined as the major dimension of response consistency, as the common factor among items, or as the unobservable ability/difficulty continuum. Latent traits identify the processes, strategies, and structures that underlie task performance. IRT assumes that a set of k latent traits underlie performance on a set of test items and that the k latent traits define a k dimensional space. Each examinee's locus in the latent space is determined by the examinee's position on each latent trait.

This paper has a twofold purpose: (1) to determine if the item responses from a test of the D and Z morphemes exhibit unidimensionality, i.e., the underlying trait measured by the test can be described in a one-dimensional latent space and (2) to characterize and to define the latent trait (s) by placing the test items at their calibrated positions along the line of the variable. If our hypothesis that the D and Z morphemes are stored as articulatory/perceptual images and that these morphemes are categorized as prototypes with their allomorphs conceptualized as a population of objects which can be considered in terms of their derivations from their prototype is confirmed, then unidimensionality must be unequivocally demonstrated.

METHOD

Subjects

This study was conducted at the Center for English as a Second Language at Southern Illinois University. Thirty-five students from the Center's level three proficiency stratum served as unpaid volunteers for the project. The distribution of subjects by native language was the following: Arabic, 7; Chinese, 7; Farsi, 1; Japanese, 1; Korean, 2; Malay, 2; Spanish, 12; Tamil, 1; Thai, 1; Turkish, 1.

At the Center, placement of students into the four full-time proficiency strata (1,2,3, and 4) is determined by the results of an Institutional TOEFL test which is administered to new students at the beginning of each term. Students who score 347 or lower are placed into level 1; students who score between 350 and 397 are placed into level 2; those who score between 400 and 447 are placed into level 3; and those who score between 450 and 497 are placed into level 4. If students score between 500 and 523 they are given additional testing (a composition and an oral interview) to determine whether they are proficient enough for full- or part-time university study or whether they need additional full- or part-time English study.

Materials

The elicitation instruments used in this research were paper-and-pencil tests, and they are presented as Figures 1 and 2. A random sample

FIGURE 1
Z Test

INSTRUCTIONS:

The possessive forms of nouns, plural forms of regular nouns, and third singular forms of verbs are all spelled with s ('s, s", es, ies), but there are actually three different pronunciations. For example, the s in the word lips is pronounced s and is referred to as /s/; the s in the word bugs is pronounced z and is referred to as /z/; the s, in the word pages is pronounced iz and is referred to as /iz/.

Pronounce each of the following words to yourself and decide how its "s" ending is pronounced -- as /s/, /z/, or /iz/. Then circle that pronunciation.

1.	bans	/s/	/z/	/iz/
2.	sees	/s/	/z/	/iz/
3.	bags	/s/	/z/	/iz/
4.	Frances	/s/	/z/	/iz/
5.	reviews	/s/	/z/	/iz/
6.	pads	/s/	/z/	/iz/
7.	shovels	/s/	/z/	/iz/
8.	shuffles	/s/	/z/	/iz/
9.	lapses	/s/	/z/	/iz/
10.	laps	/s/	/z/	/iz/
11.	beads	/s/	/z/	/iz/
12.	backs	/s/	/z/	/iz/
13.	pats	/s/	/z/	/iz/
14.	loses	/s/	/z/	/iz/
15.	tithes	/s/	/z/	/iz/
16.	banks	/s/	/z/	/iz/
17.	Fran's	/s/	/z/	/iz/
18.	choices	/s/	/z/	/iz/
19.	Pat's	/s/	/z/	/iz/
20.	prays	/s/	/z/	/iz/
21.	Rich's	/s/	/z/	/iz/
22.	clasps	/s/	/z/	/iz/
23.	Ruth's	/s/	/z/	/iz/
24.	beaches	/s/	/z/	/iz/
25.	seizes	/s/	/z/	/iz/
26.	bangs	/s/	/z/	/iz/
27.	fatigues	/s/	/z/	/iz/
28.	beets	/s/	/z/	/iz/
29.	ceases	/s/	/z/	/iz/
30.	labs	/s/	/z/	/iz/
31.	Patsy's	/s/	/z/	/iz/
32.	shoves	/s/	/z/	/iz/
33.	George's	/s/	/z/	/iz/
34.	praises	/s/	/z/	/iz/
35.	patches	/s/	/z/	/iz/
36.	taxes	/s/	/z/	/iz/
37.	Kate's	/s/	/z/	/iz/
38.	Kay's	/s/	/z/	/iz/
39.	Frank's	/s/	/z/	/iz/
40.	Anna's	/s/	/z/	/iz/
41.	wives'	/s/	/z/	/iz/
42.	badges	/s/	/z/	/iz/
43.	apes	/s/	/z/	/iz/
44.	wife's	/s/	/z/	/iz/
45	Abe's	/s/	/z/	/iz/

FIGURE 2
D Test

The past tense endings for regular verbs are spelled with ed (d, ied), but the past tense ending is pronounced three different ways. For example, the ed in the word tapped is pronounced t and is referred to as /t/; the ed in the word robbed is pronounced d and is referred to as /d/; the ed in the word nodded is pronounced id and is referred to as /id/.

Pronounce each of the following words to yourself and decide how its "ed" ending is pronounced -- as /t/, /d/, or /id/. Then circle that pronunciation.

1.	acted	/t/	/d/	/id/
2.	fixed	/t/	/d/	/id/
3.	conquered	/t/	/d/	/id/
4.	raced	/t/	/d/	/id/
5.	ached	/t/	/d/	/id/
6.	leashed	/t/	/d/	/id/
7.	boxed	/t/	/d/	/id/
8.	waited	/t/	/d/	/id/
9.	sided	/t/	/d/	/id/
10.	weighted	/t/	/d/	/id/
11.	lifted	/t/	/d/	/id/
12.	boded	/t/	/d/	/id/
13.	leased	/t/	/d/	/id/
14.	filed	/t/	/d/	/id/
15.	boated	/t/	/d/	/id/
16.	concurred	/t/	/d/	/id/
17.	lived	/t/	/d/	/id/
18.	rushed	/t/	/d/	/id/
19.	typed	/t/	/d/	/id/
20.	covered	/t/	/d/	/id/
21.	ridged	/t/	/d/	/id/
22.	filled	/t/	/d/	/id/
23.	handed	/t/	/d/	/id/
24.	bowed	/t/	/d/	/id/
25.	sighed	/t/	/d/	/id/
26.	waded	/t/	/d/	/id/
27.	rigged	/t/	/d/	/id/
28.	aced	/t/	/d/	/id/
29.	laid	/t/	/d/	/id/
30.	tipped	/t/	/d/	/id/
31.	sighted	/t/	/d/	/id/
32.	packed	/t/	/d/	/id/
33.	waxed	/t/	/d/	/id/
34.	answered	/t/	/d/	/id/
35.	ceded	/t/	/d/	/id/

of stem final sounds was selected as conditioning environments. The order of presentation was determined by a table of random numbers. Previous research has shown that these paper-and-pencil tests have concurrent validity with direct, oral tests of the D and Z morphemes (Brutten, Mouw, and Perkins, 1985).

Procedure

Both instruments were administered in the language laboratory under test conditions. Working time for both tests was 50 minutes.

ANALYSIS AND RESULTS

Table 1 presents the means, standard deviations, standard errors of measurement, and estimates of reliability for the two measures.

To answer the first research question, do the item responses exhibit unidimensionality, we submitted the data set to the Guttman-scale analysis. The Guttman implicational scaling techniques indicate whether data are truly scalable and unidimensional. If the assumption of unidimensionality is met, all the items will measure movement toward or away from the same underlying construct or trait. The degree to which a set of items exhibit unidimensionality is determined by the extent to which scores of 1 on any item are associated with scores of 1 on all items ranked as being less difficult, and inversely, a scale is unidimensional to the extent that scores of 0 on all items judged as being difficult. For a data set to be truly scalable and to exhibit unidimensionality, the obtained coefficient of scalability must be greater than 0.60.

Table 2 presents the Guttman coefficients for the D and Z morphemes. The coefficients are clearly greater than 0.60, but we have not been able to ascertain from our review of the measurement literature how many intervals above 0.60 a coefficient of scalability must be in order for a data set to demonstrate unidimensionality.

To answer the second research question, to define and to characterize the traits underlying the D and Z tests, the data were submitted to the Rasch one-parameter item response model. The Rasch model derives logits of difficulty for items and logits of ability for persons. The logits employ a true interval scale with mean difficulty set at zero. The logits

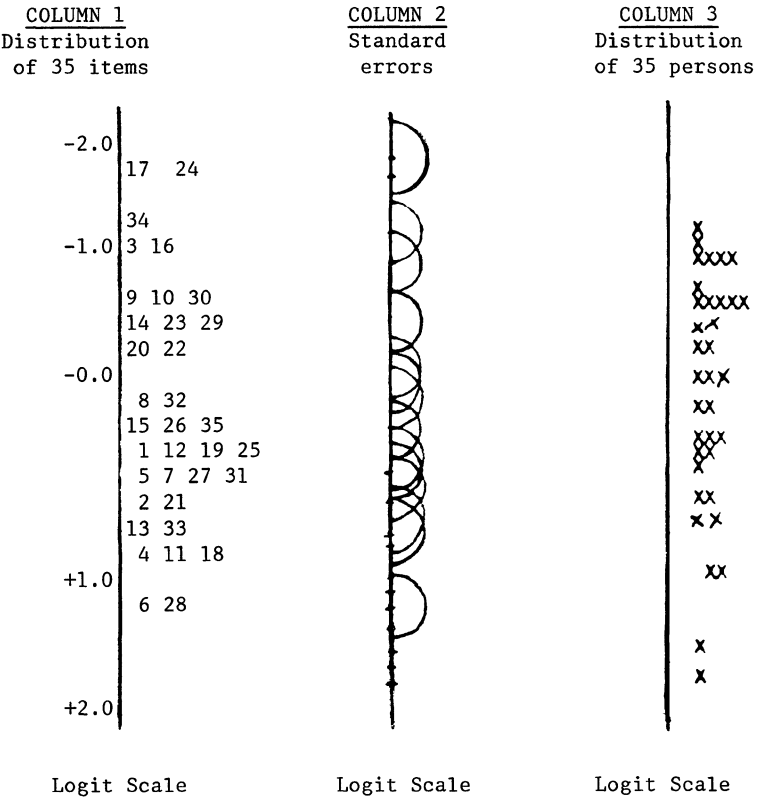
TABLE 1
Means, standard deviations, standard errors of measurement, and estimates of reliability

Test	Number of items	Number of subjects	Mean	Standard deviation	Standard error of measurement	Estimate of reliability
D	35	35	17.14	5.52	2.69	.761
Z	45	35	23.26	6.20	3.04	.760

TABLE 2
Guttman Scaling

	D	Z
Coefficient of reproducibility	.837	.839
Minimal marginal reproducibility	.321	.356
Percent improvement in reproducibility	.516	.483
Coefficient of scalability	.759	.749

FIGURE 3
Defining the D variable by item difficulty distribution, standard errors, and person ability distribution



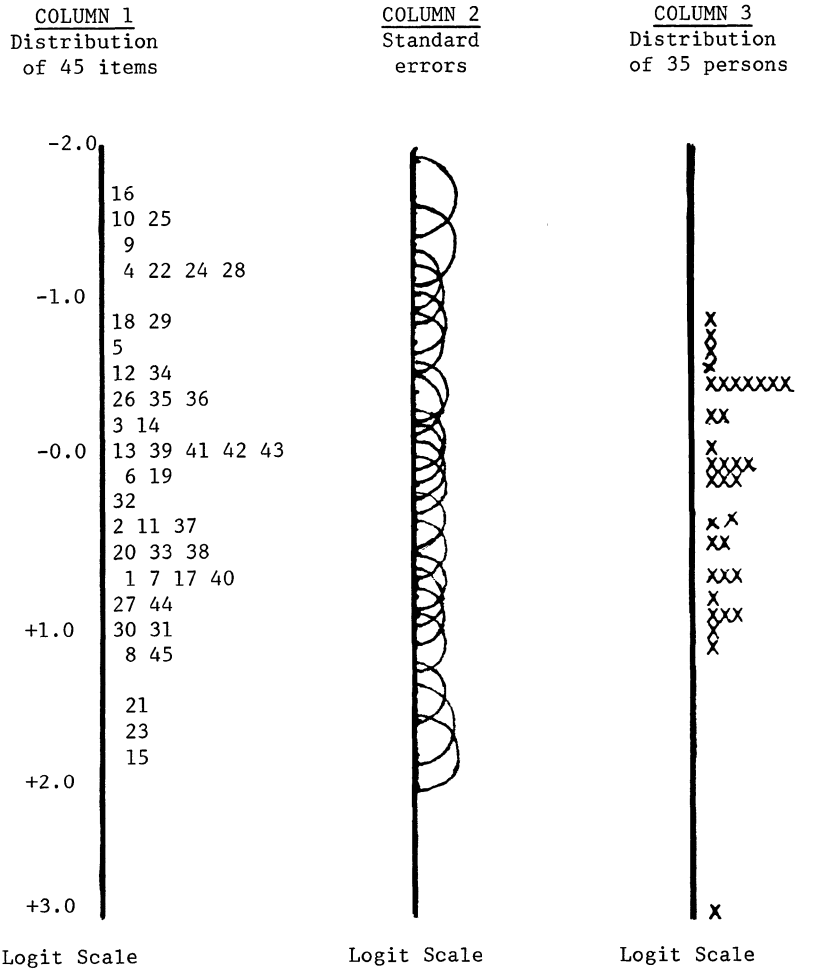
of difficulty for items are adjusted for sample spread expansion and the logits of ability for persons are adjusted for test width expansion. The negative logits represent low-difficulty items and low-ability persons, and the positive logits represent high-difficulty items and high-ability persons (Wright and Stone, 1979).

Figures 3 and 4 present Rasch descriptions of the D and Z variables. In these figures the items are placed at their calibrated positions along the line of the variable in column 1. The semicircles in column 2 are drawn to scale and represent the size of the standard error of each item difficulty. Column 3 gives the distribution along the variable of the 35 persons who participated in the calibration.

Tables 3 and 4 reflect in a more interpretable fashion the same information conveyed in Figures 3 and 4. Table 3 indicates that the easy end of the D variable is characterized by the /d/ allomorph while the difficult end is characterized by the /t/ allomorph. Table 3 also indicates that a preponderance of difficult items had word stems ending with the fricatives /s/ and /š/ and the affricate /ǰ/.

Table 4 shows that the easy end of the Z variable is characterized, in the main, by the /s/ and /iz/ allomorphs. While the difficult end of

FIGURE 4
Defining the Z variable by item difficulty distribution, standard errors, and person ability distribution



the Z variable is generally characterized by the /z/ allomorph. Like the D variable, the difficult end of the Z variable is characterized by several word stems ending in the fricatives /f/, /θ/, and /ð/ and the affricates /č/ and /ǰ/.

From these data it is apparent that ESL subjects have difficulty with words ending with fricatives and affricates. This phenomenon is similar to one of Berko's (1961) results from her study of L1 children's learning of English morphology. Berko found that children were able to form the regular noun plurals requiring /s/ and /z/, but if a word ended in /s z š ž č ǰ/, they added nothing. Another curious point is that for the D morpheme the voiceless allomorph /t/ characterized the difficult end of the variable, but for the Z morpheme, the voiced allomorph /z/ predominated the difficult end of the variable. We have expected a parallel, that is,

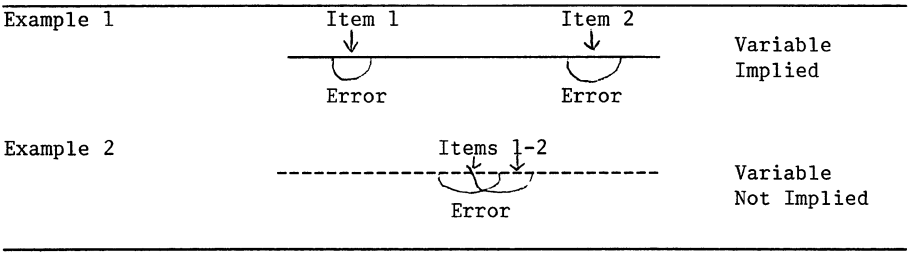
TABLE 3
D Morpheme

Logit of Difficulty	Stem final sounds + allomorphs
-1.7	v+d; ow+d
-1.2	r+d
-1.0	r+d; r+d
-0.6	d+id; t+id; p+t
-0.3	l+d; d+id; ey+d
-0.2	r+d; l+d
-0.1	t+id; k+t
+0.1	t+id; d+id; d+id
+0.3	kt+id; d+id; p+t; ay+d
+0.4	k+t; ks+t; g+d; t+id
+0.5	ks+t; ǰ+d
+0.7	s+t; ks+t
+0.8	s+t; ft+id; š+t
+1.3	š+t; s+t

TABLE 4
Z Morpheme

Logit of Difficulty	Stem final sounds + allomorphs
-1.7	k+s
-1.4	p+s; z+iz
-1.2	ps+iz
-1.1	s+iz; p+s; č+iz; t+s
-0.9	s+iz; s+iz
-0.8	yu+z
-0.5	k+s; z+iz
-0.4	ŋ+z; č+iz; ks+iz
-0.2	g+z; z+iz
-0.1	t+s; k+s; v+z; ǰ+iz
+0.1	v+z
+0.3	g+z; d+z; t+s
+0.5	ey+z; ǰ+iz; ey+z
+0.7	n+z; l+z; n+z; ə+z
+0.8	g+z; f+s
+0.9	b+z; iy+z
+1.1	gl+z; b+z
+1.4	č+iz
+1.6	θ+s
+1.8	ð+z

FIGURE 5
Defining a variable



(Wright and Stone 1979, p. 84)

either the voiced or voiceless allomorphs would have been prevalent in the difficult ends of the variables for both morphemes. We have a tentative explanation for this finding. Some of the stimulus words for the difficult end of the D morpheme do not have a high frequency of occurrence as do some of the stimulus words for the difficult end of the Z morpheme. Stimulus words for the D morpheme include ceded, boded, rigged, leashed, and aced, while the following more commonly occurring words are found in the Z morpheme list: bags, bans, wife's, Ruth's, and Rich's. We noted in an earlier section of the paper that we assume that for each of these morphemes there is a prototype and that the allomorphs can be described in terms of their deviation from the prototype. For the D morpheme we have chosen /d/ as the prototype because /t/ can be related to the prototype by a universal devoicing rule and /id/ is a separate syllable. By the same reasoning we have chosen /z/ as the prototype for the Z morpheme. Since the logits of difficulty are on an interval scale, we found the mean logit for each allomorph and then described the deviation of each allomorph from the prototype. In this fashion we found that for the D morpheme the /t/ allomorph deviates from the prototype by 0.95 logit, and /id/ deviates by 0.48 logit. For the Z morpheme the /s/ allomorph deviates from the prototype by 0.84 logit and /iz/ deviated by 0.99 logit.

We are also concerned with how we can use the Rasch analysis to generalize the definition of the variables. We do this by examining the calibrated items to ascertain what they imply about how the calibrated items disperse in a manner that shows a coherent and meaningful direction. According to Wright and Stone, the estimates of item difficulties must be well separated by several standard errors for the direction of a variable to be defined.

Figure 5 illustrates a canonical array of item difficulties showing when a direction for a variable is implied. An examination of Figures 3 and 4 indicates that for the D and Z morphemes the associated standard errors for the item difficulties overlap substantially. The net result is that the items are so close to each other that, considering their standard errors, they are not separable; consequently, we have found a point for each variable, but no direction has been established, and no variable has been clearly and unequivocally implied.

We believe that one reason for the lack of direction for the two variables is language interference. Because different language groups are represented in the subject pool, the item responses represent differential grades of facility with the different allomorphs. Thus, a collage is produced rather than a definite line. Unfortunately, the various language samples are too small to segregate and to analyze separately.

SUMMARY

We have proposed that the D and Z morphemes are stored as articulatory/perceptual images and are categorized as prototypes with their allomorphs conceptualized as a population of objects which can be considered in terms of their deviations from a prototype. We analyzed two data sets elicited from measures of the D and Z morphemes and found some evidence of unidimensionality. An item response theory analysis of the variables indicated no clear direction for either of the variables, which we attribute to language interference.

DIRECTIONS FOR FUTURE RESEARCH

Our research and our application of the Rasch one-parameter model rest on the assumption of unidimensionality in the data. In reality, of course, the assumption of unidimensionality cannot be strictly met because of a number of affective, cognitive, linguistic, and test-taking factors which impinge on test performance. These factors might include, among others, ability to work quickly, attitude, language interference, motivation, test anxiety, and testwiseness. Unidimensionality may have to be interpreted as a dominant factor or component that influences test performance. It may also be the case that a test that is multidimensional for one population is unidimensional for another population.

For our future research we will employ Lumsden's (1976) method for constructing unidimensional tests. Instead of using the Guttman-scale analysis, we will iteratively submit the data to factor analysis and remove the deviant items until a satisfactory solution is obtained. A ratio of first-factor variance to second-factor variance will be used as an index of unidimensionality.

REFERENCES

Berko, Jean. 1961. The child's learning of English morphology. In Sol Saporta (Ed.), Psycholinguistics. A Book of Readings. New York: Holt, Rhinehart and Winston.

Brutten, Sheila R., John T. Mouw, and Kyle Perkins. 1985. The effect of instruction on ESL speaker control of the D and Z morphemes. Paper presented at the Delaware Symposium VII on Language Studies. October 25.

Hambleton, Ronald K. and Hariharan Swaminathan. 1985. Item Response Theory. Principles and Applications. Boston: Kluwer. Nijhoff Publishing.

Langacker, Ronald. 1982. Space grammar, analysability, and the English passive. Language, 58, 22-80.

Lumsden, J. 1976. Test theory. In M.R. Rosenzweig and L.W. Porter (Eds.), Annual Review of Psychology. Palo Alto, CA: Annual Reviews, Inc.

Nathan, Geoffrey S. 1986. Phonemes as mental categories. In V. Nikiforidou, Mary Van Clay, M. Niepokuj, and D. Feder (Eds.), Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society (pp. 212-223). Berkeley, CA: Berkeley Linguistics Society, University of California.

Rosch, Eleanor. 1973. On the internal structure of perceptual and semantic categories. In T.E. Moore (Ed.), Cognitive Development and the Acquisition of Language. New York: Academic Press.

Rosch, Eleanor. 1977. Human categorization. In Neil Warren (Ed.), Studies in Cross-Cultural Psychology, Vol. 1. New York: Academic Press.

Rosch, Eleanor. 1978. Principles of categorization. In Eleanor Rosch and Barbara Lloyd (Eds.), Cognition and Categorization. Hillsdale, NJ: Lawrence Erlbaum Associates.

THE AFFECTIVE FILTER AND PRONUNCIATION PROFICIENCY - ATTITUDES AND VARIABLES IN SECOND LANGUAGE ACQUISITION

Robert M. Hammond

Catholic University of America
Washington, DC

INTRODUCTION

In developing his theory of second language acquisition, Krashen (1982) suggests five hypotheses: the acquisition-learning hypothesis; the natural order hypothesis; the input hypothesis; and the affective filter hypothesis. The first three of these hypotheses are central to the organization of a language program using the natural approach (Krashen and Terrell 1983); that is, they form the underlying bases for a program whose purpose it is to develop in beginning students as much communicative competency as possible in a beginning language course or series of courses. The latter two, the input and affective filter hypotheses, however, determine on a day-to-day basis what actually takes place in the second language classroom. In very general terms, the input hypothesis states that we must provide as much comprehensible input as possible for a student in the second language classroom, since within Krashen's theoretical framework, it is claimed that it is through and only through comprehensible input becoming comprehended input that language is acquired (not learned). The notion of the affective filter, originally presented in Dulay and Burt (1977), which is much less controversial, and valid for almost all language teaching methodologies, states that the affective variables of motivation, self-confidence and anxiety (Krashen 1982) have a profound influence on language acquisition (not learning). The claim of the natural approach is that students will acquire second languages best when they are in an environment which provides a maximally low (weak) affective filter, and a maximally high amount of comprehensible input.

Within the category of the affective variables, the roles of linguistic stereotyping and language learner attitudes are also significant, as these two factors may have an important influence on student progress in second language acquisition. The vast majority of adults who study second languages fail to acquire what could be described as a native or even native-like pronunciation in these second languages. Likewise, these same adult second language learners seem initially to make remarkable progress in acquiring the sound system of these second languages, only to experience a dramatic drop-off in this pronunciation process after little more than the minimally essential elements of the second language sound system have been internalized (acquired). That is, adult learners seem to achieve a certain level of pronunciation proficiency in a second language, and they then fail to significantly improve their pronunciation. As it is

hypothesized that adult learner attitudes toward second languages and second language acquisition may be important factors underlying such problems in the acquisition of a native-like pronunciation, several affective variables will be explored in the present study. Data relative to the following second language acquisition attitude variables will be presented herein: 1. the ability of the second language acquirer to accurately evaluate his/her own level of pronunciation proficiency; 2. second language acquirers' attitudes toward foreign accent; 3. the correlation of the second language acquirer's degree of pronunciation proficiency with his/her attitude toward foreign accent; 4. second language acquirers' reactions to three common linguistic stereotypes concerning foreign accent in a second language.

1. METHODOLOGY. The data under analysis in the present study come from two sources: 1. the responses given on a questionnaire administered to 282 subjects who are all native speakers of Spanish and who have studied English as a second language for at least two years; 2. a pronunciation evaluation questionnaire completed by the English instructors of these same 282 subjects. Eighty percent of the subjects are natives of Cuba, while the remaining 20% come from Mexico, Puerto Rico, Spain, Nicaragua, Chile and Colombia. Sixty percent of the subjects are female, and the 282 subjects ranged in age from 18-63 years, their median age being 27 years. Seventy-six percent of these subjects were between 20 and 49 years of age (mean age = 29.9 years; standard deviation = 13.6). Among the subjects utilized, 64% have resided in the United States for fewer than five years, and all are permanent residents of the Little Havana section of Miami, Florida. A thirty-three item questionnaire (in Spanish) was administered to each of the subjects, and the responses to only six of these items will be analyzed herein. The instructors who provided the pronunciation evaluations of the 282 subjects are all native speakers of American English. The subjects who were utilized represent 282 of the total enrollment of 288 students in 18 English as a second language classes which were taught by native speakers of American English over a two-year period. Six subjects were eliminated because they were not native speakers of Spanish.

2. SUBJECTS' SELF-EVALUATION OF AMERICAN ENGLISH PRONUNCIATION. On the questionnaire, each subject was asked to evaluate his/her own pronunciation of American English according to five different categories. Student responses to this questionnaire item are shown in the first column of Table I.

TABLE I

Overall Evaluation of Pronunciation

	Self-Evaluation %	Instructor Evaluation %
Excellent	8	17
Good	25	20
Average	49	39
Poor	17	20
Very Poor	1	4

As seen in Table I, only eight percent of the subjects evaluated their English pronunciation as excellent. Twenty-five percent of the subjects said that their English pronunciation was good, forty-nine percent felt it was average, seventeen percent said it was poor, and only one percent felt that their pronunciation could be described as very poor. When instructors evaluated the pronunciation of these same 282 subjects, column two of Table I, they rated 17% as excellent, 20% as good, 39% as average, 20% as poor and four percent as very poor. An overall comparison of the student self-evaluation and the instructor evaluations shows that they were basically in agreement. Generally speaking, those subjects with better English pronunciation tended to under-estimate the quality of their pronunciation, while those individuals evaluated as having the greatest overall pronunciation deficiency tended to over-estimate their pronunciation achievement in English.

A one-to-one correlation of the 282 subjects' self-evaluation of their English pronunciation with the evaluation of their instructors is shown in Table II. In 48% of the cases, there was a perfect correlation between subject self-evaluation and instructor evaluation, and in an additional 45% of the cases there was a close correlation (only one level of difference). Thus, in only seven percent of the cases were the subject self-evaluation and the corresponding instructor evaluation extremely different (by two or more levels). It is clear from these data that the 282 subjects were highly accurate in evaluating their own level of pronunciation achievement in American English.

3. SUBJECT ATTITUDE TOWARD ACCENTED SPEECH. Another item on the questionnaire used in the data-collection process asked the subjects to give their opinion of foreign accents in general. They were specifically asked to give their impression of foreign accents according to one of five categories. Six percent of the subjects indicated that foreign

TABLE II

Correlation - Self-Evaluation & Instructor Evaluation

Type of Correlation	Levels of Difference	Correlation Percentage
Subject Overestimate	3	1
Subject Overestimate	2	2
Subject Overestimate	1	21
Perfect Correlation	0	48
Subject Underestimate	1	24
Subject Underestimate	2	3
Subject Underestimate	3	1

accents sound very nice, 24% evaluated them as nice, 32% as bad, 9% as very bad, and 29% felt they were not important. In other words, only 41% of these subjects felt that, generally speaking, foreign accents should be considered as a negative factor in second language acquisition.

A more specific item on the questionnaire tested subject reaction to a Spanish accent in their own English pronunciation. The questionnaire statement "It is very important to try to eliminate Spanish accent in your English pronunciation." drew strong agreement in 48% of the cases, agreement in 35% of the cases, disagreement from 15% of the subjects, and a strong disagreement in only 2% of the cases. With respect specifically to Spanish accent in the English of these 282 subjects, 83% of these individuals found accent to be a negative factor. This difference between the 41% of the subjects who reported that they considered foreign accents in general as negative and the 83% reporting that a Spanish accent in their English was a negative factor is striking. However, it is not immediately obvious to what exactly such a difference is to be attributed. This difference may indicate that subjects were more tolerant of another's accent, but were far more demanding of themselves. Another possible interpretation of this wide difference is that theoretical or hypothetical situations (foreign accents in general) versus specific, real personal situations elicit very different reactions on the part of adult second language acquirers. There are, of course, other possible interpretations. The unique source of the data utilized herein, to be discussed in the conclusions to this study, may also have had a strong effect on some of the results presented.

4. CORRELATION OF SUBJECT ATTITUDE TOWARD FOREIGN ACCENT AND THEIR OWN LEVEL OF ACHIEVEMENT IN ENGLISH PRONUNCIATION. Table III presents a correlation of the attitudes expressed by the subjects toward Spanish accent in their own English pronunciation with the actual quality of English pronunciation they have achieved.

TABLE III

Correlation of Subjects' Attitude & Achievement

Subject Attitude	Subject Pronunciation Achievement				
"Elimination of Spanish accent is important."	Excellent	Good	Average	Poor	Very Poor
Strongly Agree	38%	58%	45%	61%	25%
Agree	50%	26%	40%	17%	50%
Disagree	12%	11%	13%	22%	25%
Strongly Disagree	0%	5%	2%	0%	0%
TOTALS	100%	100%	100%	100%	100%

One might hypothesize a high positive correlation between the subject's desire to eliminate foreign accent and the actual achievement of a high quality pronunciation in his/her own English. Therefore, we might expect Table III to show that the subjects evaluated as having an excellent English

pronunciation would be the most intolerant of foreign accent, and that those with the poorest English pronunciation would conversely be the most tolerant of foreign accent. The data shown in Table III, however, show inconsistent results.

Eighty-eight percent of the subjects who were evaluated as having an excellent English pronunciation expressed disapproval of a Spanish accent in their English, and 84% of the subjects whose English pronunciation was evaluated as good also expressed disapproval of Spanish accent in their English. An analysis of the data for only these two sub-groups of subjects yields an expected high positive correlation between Spanish accent disapproval and the quality of English pronunciation achieved by the subjects. Up to this point, a hypothesis predicting high pronunciation achievement and lack of foreign accent tolerance appears to be tenable. Unfortunately, the data from the other three sub-groups of subjects who had the lowest levels of English language pronunciation achievement do not support such a hypothesis. The data in Table III indicate that all five sub-groups favor the elimination of Spanish accents in English. Specifically, 85% of the subjects with an average English pronunciation, 78% with a poor pronunciation, and 75% with very poor pronunciation also expressed disapproval of Spanish-accented English. Based on an analysis of these data, there seems to be no obvious correlation between tolerance of foreign accent and quality of second language pronunciation. There does seem to be some movement toward a positive correlation in these data, in that the percentages of subjects demonstrating excellent English pronunciation to very poor pronunciation do move generally through the five sub-groups from higher to lower percentages of disapproval of foreign accent, 88%, -84% -78% -75% respectively. These observed differences in percentages, however, are not statistically significant, at least not based on a data sample of this size.

4. SPEAKER INTELLIGENCE AND FOREIGN ACCENT. The questionnaire also included an item which sought subject opinion concerning the supposed relationship between language learner effort and intelligence and the ability to eliminate foreign accent. Subjects reacted as follows to the statement "Hard-working and intelligent people can always succeed in eliminating foreign accent.": 19% of the subjects indicated strong agreement, 32% agreement; 37% disagreement; 8% strong disagreement; and 4% of the subjects didn't know if the statement was true or false. That is, 51% of these subjects believed the stereotype that if you are intelligent and make an effort you can always rid yourself of a foreign accent. Forty-five percent of the subjects expressed disagreement with the supposed relationship expressed in this questionnaire statement.

6. JOB SUCCESS AND FOREIGN ACCENT. The 282 subjects employed in the present study were also asked how they felt about the relationship between employment success and foreign accent. The questionnaire statement read "A foreign accent decreases an individual's probability of job success." Thirteen percent of the subjects strongly agreed with this statement, 28% agreed, 40% disagreed, 14% strongly disagreed, while 5% did not know. Therefore, most of the subjects, 54%, did not believe the stereotype that a foreign accent had a negative effect on probability of job success. It is interesting to note that although 83% of these same subjects believed that a Spanish accent in their English was a negative factor, only 41% of them believed that foreign accent in general was detrimental to job success. Although these two findings seem to contradict one another, there is an explanation for this seeming discrepancy due to the nature of the data source. As mentioned previously, this matter will be discussed in the conclusions to this study.

7. SECOND LANGUAGE COMMUNICATION AND FOREIGN ACCENT. In the final questionnaire item included in the present study, subjects were asked to

agree or disagree with the statement "If an individual can express himself in a second language, the fact that he speaks with a foreign accent is of little importance." Thirty-three percent of the subjects strongly agreed with this statement, 44% agreed, 15% disagreed, 5% strongly disagreed and 3% didn't know. Once again, although 83% of these individuals believed that a Spanish accent in their English was a negative factor, only 20% of these same subjects felt that the presence of a foreign accent was detrimental if an individual could otherwise express himself in a second language.

8. CONCLUSIONS. Before presenting any conclusions based on the information presented in the present study, it should be stated that the data analyzed herein must be considered tentative due both to the limited size of the corpus under analysis (282 subjects) and the interim nature of this study. In the near future a great deal more data that has now been gathered will be analyzed from a more diversified group of subjects. Furthermore, the remaining 27 items from the subjects' questionnaire will be analyzed and correlated with the data just presented. With these limitations in mind, these data would seem to support the following conclusions:

1. Most of the second language learners queried believe that if you are intelligent and make an effort at it, you can always rid yourself of a foreign accent. They also felt that foreign accent was <u>not</u> a detriment to job success.

2. Only 20% of these same subjects felt that the presence of a foreign accent was important if a speaker could otherwise communicate in a second language.

3. Second language learners appear to be highly capable of accurately evaluating the quality of their own pronunciation in a second language.

4. Most of the subjects utilized herein (59%) did not object to foreign accents in general, but the great majority of them did object to the presence of a <u>Spanish</u> accent in their English pronunciation.

5. There appears to be no apparent correlation between a second language learner's disapproval of foreign accent and the actual acquisition of a high-quality pronunciation in a second language.

As previously mentioned, it is hypothesized that some of the apparent contradictions present in the data utilized in the present study may be explanable in terms of the unique sociolinguistic nature of the bilingual community in which all 282 subjects reside.

The so-called Cuban section of Miami, La Sagüesera, is generally understood by area residents to include most of the southwest quadrant of the city of Miami, a large portion of Miami's northwest quadrant nearest Flagler Street, a large portion of the city of Coral Gables, and the cities of Westchester and Sweetwater. This is an area of more than 400,000 persons of Cuban origin (United States Bureau of the Census 1982:134) and of a total of more than 750,000 Hispanics (MacDonald, 1985:45). In this loosely defined area, it is extremely easy to acquire <u>any</u> goods, services or other needs in Spanish. There are Spanish-speaking hospitals, police stations, grocery stores, restaurants, dentists, florists, schools, funeral homes and every other possible business or service institution one could ever need from birth to the grave. Therefore, the actual need to speak English to be able to live in Little Havana is only minimal. This fact may have had a direct influence on the opinions expressed by the 282 subjects.

Another important factor bearing on the data in this study is the

ready availability of jobs to monolingual Spanish speakers in Little Havana. Once again, jobs of almost every nature are available in this area, and the need to speak English well can be of little consequence to job success. While speaking English well in Miami can provide Hispanics with more potential jobs and potentially greater opportunities for upward social and employment mobility, speaking English well is not a requirement for either survival or job success.

Also, many Americans in the Miami/South Florida are either bilingual in English and Spanish, functional (in varying degrees) in Spanish, or at least skillful in understanding Spanish accented English. Unlike the case in most immigrant situations in which the immigrant has had to learn to cope linguistically and culturally with his new environment, in Miami, due largely to the economic opportunities presented by the large Hispanic community, many Americans have had to learn at least some Spanish to survive. This situation may have strongly influenced the subjects' feelings concerning the relative importance of accent if communication is achieved.

To test the validity of the above hypothesis concerning the uniqueness of the Miami/South Florida linguistic community, several follow-up studies to the present one could be made. For example, the questionnaire used herein could be given to Hispanics in the United States who live in areas other than Miami. A comparison of these two bodies of data would prove interesting. Also, the same questionnaire given to Hispanics who live in other United States communities which have large Hispanic centers, such as Los Angeles, New York City, and San Antonio, would provide further data. One further interesting possibility would be to administer an English version of the questionnaire used in the present study to English speakers who have daily contact with Hispanics in areas such as Miami. Hopefully, the present study has provided some useful data concerning attitudes and stereotypes about second language learning. It is clear, however, that this is an area in the field of second language acquisition that needs to be investigated in much greater detail before any reliable conclusions can be drawn concerning the affective variables discussed in this study.

BIBLIOGRAPHY

Dulay, H. and M. Burt. 1977. Remarks on creativity in language acquisition. M. Burt, H. Dulay and M. Finnochiaro (Eds.). Viewpoints on English as a second language. New York: Regents Press, pp. 95-126.

Krashen S. 1982. Principles and practice in second language acquisition. New York: Pergamon Press.

Krashen, S. and T. Terrell. 1983. The natural approach. Hayward, CA: Alemany Press.

Levitan, A. 1980. Hispanics in Dade County: their characteristics and needs. Miami, FL: Office of the County Manager.

MacDonald, M. 1985. Cuban American English: the second generation in Miami. Ph.D. dissertation, Gainesville: University of Florida.

United States Bureau of the Census. 1982. General population characteristics, Florida. Washington, D.C.,: United States Government Printing Office.

PROFICIENCY WITHOUT SUPPORT:

"MY PARENTS NEVER SPOKE IT TO ME"

Marion Lois Huffines

Bucknell University
Lewisburg, PA

In the United States, the issue of language proficiency is raised primarily in the classroom where native speakers of American English study foreign languages to achieve their personal and career aspirations and where recent immigrants learn English to facilitate their passage into American society. Assessment of proficiency in this context is necessary to mark progress toward fluency. Another albeit small group of American native English speakers also strive toward other language proficiency far removed from the American classroom.

Descendants of immigrants to the United States whose families have maintained a minority language up to a previous generation often strive to learn the minority language which their families no longer regularly use. They do so without the support of family, school, and a linguistic hinterland. Indeed, these speakers achieve limited proficiency in spite of resistance offered by the family and school authorities.

As minority language speakers adopt English to fulfill more and more of their communicative needs, children have ever decreasing access to the language of their parents. These learners must often resort to using grammatical forms and patterns from English in the effort to produce the minority language; but an explanation of structures used by these nonfluent speakers which is based on structural convergence to English is not sufficient to account for the loss of some features and the retention of others in their variety of the receding language. See especially Dorian (1981), Tsitsipis (1981), and Schmidt (1985).

Most varieties of American German are in various stages of decline. Only among the separatist religious groups does American German enjoy the continued support of native speakers. This study reports the efforts of nonsectarian Pennsylvania Germans to learn Pennsylvania German without the usual cultural supports. Pennsylvania Germans who are the first generation in their respective families to learn English natively achieve a proficiency in Pennsylvania German sufficient to fulfill a few communicative functions. Convergence to English is often given as the explanation for the non-German features found in their speech. English is cited as the ultimate culprit for the loss of inflections, the trend toward main clause word order in subordinate clauses, the preferential use of haben to form all perfect tenses and sein to form the passive, and the increased use of tun as an auxiliary (Enninger 1980, Costello 1978).

The assumption that Pennsylvania German is converging toward English as it dies is seldom challenged, and a frequently cited manifestation of that convergence is the loss of the dative case. The role which English plays in the loss of the dative in other varieties of American German has been argued variously: some scholars suggest that the loss of the dative/accusative distinction is purely an internal development, a result of a pattern inherited from the parent European dialects (Gilbert 1965). Others argue that the loss is almost entirely due to the influence of American English (Eikel 1949). While I can produce data from my own research as can others which demonstrate that younger speakers use the dative far less frequently than do older speakers, I propose to argue that the dative case is not being lost, but that the language itself is dying. Nonfluent speakers formulate systematic rules for dative usage. These rules reflect limited exposure to the language, but are based nevertheless on Pennsylvania German structures.

Procedures

The following observations are based on interviews with 33 Pennsylvania Germans who live in central Pennsylvania. These speakers are divided into three groups: 13 native speakers of Pennsylvania German (Group N), 9 first in the family native English speakers (Group 1), and 11 second or later in the family native English speakers (Group 2). The interview consisted of three parts: free conversation, translation of English sentences into Pennsylvania German, and picture descriptions. I spoke English during the interview; the informants responded in Pennsylvania German, a common discourse pattern in communities where a minority language is receding (cf. Dorian 1981). The results of the translation task will be reported separately from those of the free conversation and picture descriptions.

Results

In order to assess the path of nonfluent speakers' proficiency in Pennsylvania German and the extent to which their Pennsylvania German converges toward American English, I consider here the use of the dative case in one of its functions, its use to express possession. The expression of possession is an area in which one can evaluate both acquisition and linguistic convergence in spite of the close genetic relationship between English and Pennsylvania German.

In Pennsylvania German possessive constructions, the possessor occurs in the dative and is followed by a possessive adjective which agrees with the possessed, the noun it modifies. For example, the sentence "Where is Daddy's book?" is rendered wu Is əm dadi sai bUx; the English form "whose" is rendered by the dative form of the interrogative pronoun plus the masculine/neuter possessive adjective, giving vEm sai; for example, vEm sai hUnd Is sEl "whose dog is that". The common case in Pennsylvania German fulfills nominative and accusative functions.

In the translation task, 6 sentences elicit expressions of possession. (See Table 1.)

Table 1
Case of the Possessor
(Translation Task)

Group	dat	common	-s	other
N	25	21	2	3
1	19	17	6	2
2	9	10	18	12

Native speakers (Group N) respond most frequently with dative forms to express the possessor. The two youngest members of Group N account for 6 of the 21 common case forms. These two speakers have usage profiles which parallel those of speakers in Group 1. The native English speakers (Groups 1 and 2) show considerable variation. Group 2 resorts most frequently to the English genitive -s and also produces the most aberrant forms and other faulty constructions lacking grammatical agreement.

Free conversation and picture descriptions allow for the use of Pennsylvania German without overt reference to English. Although expressions of possession are not frequent, these data reproduce the general trends found in Table 1 more dramatically. (See Table 2.)

Table 2
Case of the Possessor
(Free Conversation and Picture Descriptions)

Group	dat	common	-s	other
N	4	3	0	0
1	3	4	0	0
2	0	0	2	4

In Group N, common case forms are given only by the two youngest speakers. Group 2 demonstrates no mastery of the possessive construction. Of interest is the category labeled "other"; here again we find forms which are not correct formulations of either the dative or the common case.

A closer look at this "other" category in the translation data reveals a progression of constructions which demonstrates speakers' efforts to produce acceptable Pennsylvania German possessives. List 1 is a representative sample showing the possessive formulations in the two sentences which translate "Whose dog is that?" and "We were in my aunt's house yesterday."

List 1: Constructions Expressing Possession

1. dative marking on possessors (the correct PG possessive)

   vEm sai hUnd Is sEl
   In mainrə ænt irəm haUs

2. incorrect dative marking

   vEm saim hUnd Is sEl
   vEm hat sElə hUnd
   In maim ænt irə haUs

3. no case marking (common case)

   var sai hUnd Is sEl
   In mai ænt irə haUs

4. possessive adjective and word order maintained

   In mai ænt sai haUs
   In mai ænts irə haUs

5. English genitive -s (often with dative marking)

   vars hUnd Is sEl
   In maim ænts haUs

In mai ænts haUs
In ǝm dadi sai šwEšdǝrs blats

Even while using the English genitive -s, speakers' efforts to employ the dative case are apparent. The last example is a case in point: the speaker uses a correct and probably formulaic dative form to translate "aunt" ǝm dadi sai šwEšdǝr, having already identified [ænt] as English, but the speaker capitulates to the English genitive -s on šwEšdǝrs in the syntactically difficult construction he creates.

As noted above, during the free conversation and picture descriptions, Group N produces mostly dative case forms, and Group 1 mostly common case forms to express the possessor. Group 2 speakers produce a variety of constructions in which they seemingly try to employ the dative. See List 2.

List 2: Group 2 Expressions of Possession

1. Ic hap dǝ hInkǝl sai ɔiǝr griyǝ mUsǝ
   "I had to get the chickens' eggs"
   (a possible dative formulation but an incorrect possessive adjective)

2. Ic hap del fǝn mai grosmamis / grosmami sai rEsIpis
   "I have some of my grandmother's recipes"

3. Un Ic fIn en / eni Uf maiǝ dɔxtǝr irǝ kɔp
   "and I find one on my daughter's head"

4. di aldǝ man hat di buƀǝ Un di buƀǝ sai mami senǝ
   "The old man saw the boy and the boy's mother"
   (word order and the possessive adjective are maintained; sentences 2 and 3 show repairs)

5. Ic war gǝbɔrǝ In mai mamis Un dadis hemEt
   "I was born in my mother's and father's house"
   (use of the English possessive ending)

It is remarkable how seldom the nonfluent speakers of Group 2 actually resort to the English construction when not forced to express possession as in the translation task. In spite of other errors in grammatical agreement and semantic appropriateness, these speakers are for the most part using some version of the Pennsylvania German grammar rule.

Discussion

For nonfluent Pennsylvania German speakers, the data do not present so much a pattern of convergence to English as a pattern of loss of a community norm. Native speakers have retained the dative case; only the two youngest in Group N make extensive use of the common case. Group 1 uses the common case but also demonstrates its mastery of dative forms. Speakers in Groups N and 1 seldom resort to the English genitive -s and do so only during the translation task. Group 2 differs from Groups N and 1 not only in the frequency of English -s forms but also in the frequency of faulty Pennsylvania German forms. These faulty constructions represent misfired strategies at maintaining discourse in Pennsylvania German. The progressive loss of the Pennsylvania German norm does not immediately result in English forms. Faulty case marking first replaces correctly formulated dative forms. Common case forms occur seemingly by default when case marking is absent. In the absence of case marking word order is maintained with the requisite number of lexical slots. When the English genitive -s first appears, it often does so within a construction which has dative marking. The English possessive does not replace all other strategies for forming

the Pennsylvania German possessive. Other strategies continue to operate in spite of increasing Pennsylvania German influence.

In this Pennsylvania German community, a dative/common case merger will never be completed: the language will die before the case merger occurs. With this generation the transmission of Pennsylvania German ends. The strategies which nonfluent speakers use to speak Pennsylvania German are remarkably free of reliance on English rules and forms. Their Pennsylvania German shows evidence of a creative (though often faulty), systematic application of Pennsylvania German rules and rule manipulation, not the superimposition of English rules on Pennsylvania German discourse. Proficiency without support, that is, without regular interaction with native speakers in contexts where the minority language must fullfill communicative functions, is possible in spite of the minimal input. Proficiency manifests itself in the application of rules producing systematic linguistic formulations with little reliance on native language structures. This proficiency without support relies on speakers' determination to create acquisitional opportunities, and their success in creating such opportunities will determine the upward limits of their proficiency.

REFERENCES

Costello, J.R. 1978. Syntactic change and second language acquisition: The case for Pennsylvania German. Linguistics 213: 29-50.

Dorian, N.C. 1981. Language death: the life cycle of a Scottish Gaelic dialect. Philadelphia: University of Pennsylvania Press.

Eikel, F. 1949. The use of cases in New Braunfels German. American Speech 24: 278-281.

Enninger, W. 1980. Syntactic convergence in a stable triglossia plus trilingualism situation in Kent County, Delaware, USA. In H.P. Nelde (ed.) Sprachkontakt und Sprachkonflikt, Zeitschrift für Dialektologie und Linguistik, Beiheft 32. Wiesbaden: Franz Steiner.

Gilbert, G.G. 1965. Dative versus accusative in the dialects of central Texas. Zeitschrift für Mundartforschung 32: 288-296.

Schmidt, A. 1985. Young people's Dyirbal. Cambridge: Cambridge University Press.

Tsitsipis, L.D. 1981. Language change and language death in Albanian speech communities in Greece: a sociolinguistic study. Ph.D. Dissertation: University of Wisconsin-Madison.

PROFICIENCY IN TECHNICAL ENGLISH:

LESSONS FROM THE BRIDGE

Lilith M. Haynes

Fachbereich 3: Literatur- und Sprachwissenschaften
Universität Gesamthochschule Essen and Carlton College
Bundesrepublik Deutschland Northfield, MN

"Dear Naomi," the note began, "What IS a bridge?" We had pored over maps of the area a fortnight before, trying to find a way to get to Newark from Washington, D.C. without crossing a bridge, and a week later I had called to report that when the Havre de Grâce exit had appeared, I had only too gleefully fled from the jammed Interstate and crossed the bridge before I knew I had come to it, not even using the left lane, which this time around was reserved for drivers with decals. And now, another week and four hours over icy Pennsylvania roads later, I was beginning to smile at the warnings preceding these things that distinguished themselves only by railings approximately as high as my slush-spattered Escort. "But," the note went on, "the Catawissa did give me pause!"

The Catawissa? There, nestling in the Pennsylvania hills, cool and straight and flat, its pristine white sides topped at intervals with so many US flags snapping to attention in the crisp midday wind, long and narrow over the curving icy mass for which it is named, lay the Catawissa: a bridge. Oh no, not after all I have gone through getting here! And, on the way home, could this be the same thing? Curving, wet, flags limp and bedraggled, indistinguishable, almost, from their poles; the bridge looking out to the mist and smoky hills I would have to traverse to make the car-pool at 2:55? Bridges. Bridges. Bridges? Every Guyanese has to cross them -- Guyana means "land of many waters", for heaven's sake. I had had to cross one every time I went from the house to the street, with a bicycle, on foot, carrying packages. Then, the bridges over the Constitution River and the Careenage in Barbados, and, in the ensuing decades, bridges over the San Francisco Bay, New York's waterways, the Rio Grande, the Nile, the Mississippi, the Ruhr, the Rhine, crossed to earn my daily bread. What IS a bridge? It is, I tell myself, a way to get to there from here. In between, there is no Festland: hence, a bridge. And, as I write, I recall the idiocies of translators who have tried to tackle my German job-title of Brückenkursleiterin Englisch. What? I teach bridge courses: design them, decide when a section is too small to teach (two students) or large enough to warrant splitting into two (over thirty), determine when they meet, set the examinations, choose their dates, grade them, determine who passes or fails. I should know about bridges. What IS a bridge? What are lessons from the bridge?

A bridge course, in the newer German university systems, is a course charged with the responsibility of providing a student who has not completed

the Abitur (the traditional secondary-school certification for university matriculation) with the skills to perform in the chosen field of university study on a par with holders of the Abitur: a way to get to there from here, a going over and out, a straight way, with supports, as direct and short as possible. At their inception, there was political opposition to bridge courses, but I came to them after the students had succumbed to doing English, German, History and Mathematics bridge courses, trying to get quickly past them and on to the more interesting subjects which they had enrolled to pursue. I came to bridge courses after years of plying the various trades of the applied sociolinguist. I could speak very little German, I could read a little more, and I could write much less. My students had studied English for periods ranging from one to eight years and, for the most part, had not used the language for similar spans of time. I was to teach them to read and write in English about topics in their fields of proposed concentration -- Chemistry, Physics, Graphic and Industrial Design, Civil and Mechanical Engineering. I had done little science beyond "O" levels two decades before, and I couldn't even drive a car! The Ministry of Education in North Rhine-Westphalia provided some guidelines: the courses will meet for two hours per week for two semesters -- i.e. almost forty eight hours in all; the examination will consist of replying in English over a four-hour period to four questions based on a text of 500 words; a student who fails an examination three times will not be permitted to complete the first part of the diploma. What IS a bridge?

Right at the beginning, running alongside my apartment building, there was the Ruhr, and so, a bridge -- to the local train, or, leaving the driving to the Deutsche Bundesbahn, a long-distance bus, which allowed me forty minutes of vocabulary study in the Oxford-Duden (a marvellous collection of English and German technical jargon arranged by topic and accompanied by drawings and diagrams) to and from work. The texts would be the crux of the course, and I started off with pablum, copying passages from children's encyclopedias and news magazines, training myself in the technical content before moving on to professional journals and textbooks in these fields. In a technical English class, the goals are specific; not only must the technical be transmitted, so must the English. Here, thought everyone, the road is clear. Applied sociolinguists know, after all, that total immersion in the target language is the way. So, I speak English. I try it slowly, even, but it doesn't work! They understand too little. They say that my accent is not British, and that's what they are accustomed to. Never mind that most of this reflects BFBS varieties (British Forces Broadcasting Service); I do a quick survey of World English, German dialectology, and bridge crossing. We conclude that I do know more German than they do English, and agree that I will use German to teach. I warn them that since they have all had more English language instruction than I have had in German, I will only accept criticism of my German in English.

The two semesters prescribed by the Ministry cannot in fact be used, since departments schedule time for bridge courses only in the first semester of a freshman's program, and so the bridge course becomes a condensed, if not intensive, attempt to teach in eleven weeks what eleven years have not done. What IS a bridge course then? It is, it must be, a way that permits no stopping, standing or waiting in the comprehension of the content and composition of a text. So the text is supreme. In this affluent society, students are given xerox copies of text selections once the topics of interest are identified and agreed upon: a little book, so to speak, of articles on related technical topics. With each text, the first task is to read and understand its content, with some attention paid to acceptable pronunciation, but emphasis placed on the composition of the text. Like a bridge, a text always has a beginning, a middle, and an end: so must the students' writing also be. The middle might describe an ex-

periment, it might provide a taxonomy and explanation of a variety of processes or products; it might compare two experiments or procedures; it might evaluate a number of competing issues; but it cannot start without a beginning and stop without an end. The beginning must introduce and prepare; the end must make findings and come to a conclusion. This is tough work, because the technical environment epitomises knowledge of the middle: is it going to be a suspension bridge; what kinds of cables and riveting will deal with the different forces effectively; what is the optimum span; would a tunnel be cheaper? But the approaches are vital, and the students have to become as adept at evaluating and establishing the composition and subsistency of the soil; the nature, volume and direction of the traffic coming and going there from here, as they must be in selecting oil- or water-hardened steels for different applications.

Most textbooks for technical English courses I found inappropriately simple in their technical content, and so longwinded in their analysis that they rarely got past teaching the composition of individual sentences about this content. In bridge courses, however, the students have to write, after eleven weeks, for four hours, and so the "course-leader" must concentrate upon connected prose. And so to syntax. Here again, thought everyone, the road is clear. Applied sociolinguists know, after all, what English grammar is all about: contrastive analysis, error analysis, universals of language-teaching, they have all shown which features will be problematic in different language contact situations; everyone knows that third-person singular markers, plural-s, apostrophe-'s will be the potholes and obstacles to written proficiency. But here, cool and secure like the Catawissa, nestled in the history of English syntax, lying in wait for the German student wanting to get there as quickly as possible from here, sit bei and a few other prepositional cousins, the predilection of the technical register for the passive, and, like the precipitation which makes for skids when the road is not yet slippery, the sequence of tenses (and particularly two difficult points of contact, the perfect and the simple present). Oh no, not this, not now. What IS a bridge, after all? How to manoeuvre proficiently, without pause? Radio on, sing, stay in the left lane, look straight ahead, over? Maybe. Better to concentrate on the facts, though, that the design is coupled, at least, with the function, that content exists only with form, that practice makes for perfection. Most of class-time at the beginning is devoted to reviewing the tense-forms and aspectual markers, drilling them in writing, and testing them with cloze procedures. As the formulae and rationales are mastered, homework assignments transfer their employment to essays based on the texts, and towards the end of the semester, to timed essay-writing in preparation for the examination.

Fortunately, and most often in the most technical subjects, there is always a class clown who helps to lighten the atmosphere of doggedly pressing ahead to the other side, the Schein, the certificate of proficiency, because, more often than not, the student in the bridge course does not receive this signet on completion of the examination. A certificate of proficiency not only acknowledges that its holder has completed the required hours of study and shown ability to apply the knowledge thus acquired in normal circumstances; it has to deal with the Catawissa, too, lying unexpectedly around the next bend in the road, icy deck and all. Since English-department faculty teach only two eleven-week sessions per year, and are expected to hold office-hours at least once per week when not on the official 30-day vacation days allotted to all civic servants, my way of providing the driving practice, the buttresses to proficiency, honor the tradition of the rewrite in this non-teaching period. Right at the beginning, I hang on my doorknob a clipping which few students fail to read:

"[Henry Kissinger] used to tell a story about a Harvard professor

who demanded that a student write ten successive redrafts of a term paper, returning each one with the query, "Can't you do better than that?' After the tenth draft the student exclaimed in exasperation: "No, I can't do better than that." "In that case," the professor said, "now I will read it."

I assure bridge course students, and the red markings attest, that I do read every draft of their answers to the final examination. For a few students, the examination is indeed final, and they receive their Schein; for just as few, when comprehension of the text is so dismal as to forbid written responses of any length, the course must be repeated and a new examination written: they fall off the bridge, into the icy morass of rescheduling their time; for the great majority, there is the need for two rewrites under controlled honor-conditions (whose violations are easily spotted, and penalized), and sometimes three.

When they collect the Schein, the students always leave with an auf wiedersehen, which I elaborate to let them know that I mean this quite literally, that keeping the proficiency acquired for the bridge means coming to advanced seminars, reading, writing, turning on the radio to, singing, and talking, English. After eight years on the bridge, though, I still ask, "What IS a bridge?" For every set of students is -- a Catawissa.

REFERENCES

Angelis, Paul J. 1975. "Sentence combining, error analysis, and the teaching of writing". Burt, Marina K. and Heidi Dulay (Eds.) On Tesol 1975, Washington, D.C.: TESOL

Christ, H. and H.E. Piepho (Eds) 1977. Kongressdokumentation der 7 Arbeitstagung der Fremdsprachendidaktiker, Limburg, Franconia.

Corder S. Pit. 1975. "Error Analysis, interlanguage, and second language acquisition". Language Teaching and Linguistics 8. 201-217.

Dudenredaktion and German Section of Oxford University Press, Dictionary Department, Oxford-Duden Bildwörterbuch Deutsch und English, Dudenverlag 1979, Bibliographisches Institut Mannheim.

Duskova, Libuse. 1969. "On sources of errors in foreign languages" IRAL 7: 11-36.

Hakuta, Kenji. 1986. Mirror of Language: The Debate on Bilingualism Basic Books. New York.

Kissinger, Henry A. White House Years, Little, Brown, & Co., Boston 1979.

Mukattash, Lewis M. 1981. Error Analysis of Jordanian Students' English, University of Jordan Monograph.

Richards, Jack C. 1979. Introducing the perfect: An exercise in pedagogic grammar. TESOL Quarterly 13.4: 495-500.

Ulibarri, D.M. 1982. Limited-English proficient students: A review of national estimates. Report TN-10 November. Los Alamitos, California: National Center for Bilingual Research.

Weinreich, Uriel. 1953. Languages in Contact: Findings and Problems. New York, Linguistic Circle of New York. Reprinted, Mouton The Hague 1974.

Wode, Henning, 1979. "Operating principles and 'universals' in L1, L2 and FLT" IRAL 17.3: 217-231.

LANGUAGE PROFICIENCY IN READING AND WRITING

TRAINING FORMAL SCHEMATA - REPLICATION RESULTS

Patricia L. Carrell

University of Akron
Akron, OH

INTRODUCTION

A number of research studies have empirically shown that the rhetorical organization of a text interacts with the reader's formal schemata -- i.e., the reader's background knowledge of and experience with textual organization -- to affect reading comprehension. This effect of text structure on reading comprehension has been shown to be operative for both narrative and expository texts. For example, the work of Thorndyke (1977), Mandler (1978, Mandler and Johnson 1977, Johnson and Mandler 1980), Rumelhart (1975, 1977), and Kintsch (1974, Kintsch and van Dijk 1978) has shown that different patterns of rhetorical organization of English narrative prose affect the way prose is understood and recalled by native speakers of English. The work of Meyer and her colleagues (1975, 1977a, 1977b, Meyer, Brandt, and Bluth 1980, Meyer and Freedle 1984) has shown similar effects on native-speaker comprehension of English expository prose. Furthermore, these effects on reading have been demonstrated via differing measures of comprehension -- written recall protocols, summaries, retellings, and question-answering. Since the latter research on expository prose has provided further evidence that knowledge and use of textual organization -- specifically what Meyer calls the "top-level" organization -- discriminates good readers from poor readers (Meyer, Brandt, and Bluth 1980), it is reasonable to ask whether instruction which focuses on text structure improves comprehension for poor comprehenders. Several recent studies [Bartlett 1978, Gordon 1980, Short 1982, Singer and Donlan 1982, Geva 1983, Mosenthal 1984, Taylor and Beach 1984, Reutzel 1985] have found that teaching various aspects of text structure can improve comprehension for readers of English as a native language.

In the domain of training on narrative prose, three different studies have been reported -- two dissertations, one by Christine Gordon, one by Elizabeth Short, and an article by Singer and Donlan.

In the domain of training studies with expository text, there have been at least five studies. Again, due to time constraints, these will not be reviewed in depth. Suffice it to say that Geva's study involved teaching a text-mapping strategy called 'flowcharting,' and that Reutzel's

study also used visual story maps and encompassed both simple narrative and various types of expository texts. (Text-mapping involves selecting key content from a passage and representing it in some sort of visual display -- boxes, circles, connecting lines, tree diagrams, flow charts -- in which the relationships among the key ideas are made explicit, e.g., topic, elaboration, cause-effect, process, example, detail, conclusion. Geva's flowcharting is one type of text-mapping strategy.) The Bartlett (1978) study is the one which provided the framework for this training study which will be described in more detail shortly. All of these training studies show that reading comprehension can be significantly facilitated by explicitly teaching readers about the rhetorical organization of texts (narrative or expository) and by teaching various strategies for identifying and utilizing that structure during the reading process.

In second language reading, relevant research is scarce. Some recent research has begun to investigate the effects of rhetorical organization on second language reading comprehension. Carrell (1984a) has shown the effects of narrative rhetorical organization on ESL reading comprehension. Connor (1984, Connor and McCagg 1983a, 1983b) has examined the effects of one type of English expository organization on Japanese and Spanish ESL readers. Carrell (1984b) has shown the differing effects of four types of expository organization on ESL readers of Arabic, Spanish, and Oriental backgrounds. Hinds (1983a, 1983b) has shown differing effects of a traditional Japanese organization on Japanese readers and native English readers.

In a recent article, Carrell (1985) showed that facilitating effects on second language reading could be obtained by explicit teaching of text structure. To date, however, that study remains the first and only one of its kind in second language reading. This paper reports a replication of the original Carrell (1985) study. Because the Carrell [1985] article is readily available in the TESOL Quarterly, and because it contains an extensive discussion of the literature just reviewed, any of you who are interested may refer to the article for further discussion of this previous research. We shall, therefore, move now to the study itself and the most recent results, with a larger N-size, combining the earlier results with an N=25 with the new results with an N=20. The joint results with the new N=45 are even more robust and give even stronger indication that reading in English as a second language can be facilitated by explicitly teaching aspects of expository text structure.

This Training Study

Subjects. This study was conducted with a heterogeneous group of 45 high-intermediate proficiency ESL students, Level 4, enrolled in the intensive English program for foreign students (CESL) at Southern Illinois University at Carbondale. TOEFL scores of students placed in Level 4 generally fall in the range of 450 to 500. The native languages represented included: Chinese-11, Spanish 9, Bahasa Malaysian-7, Arabic-6, Japanese-4, Indonesian-3, Korean-3, Turkish-2.

Training Procedures. Training was based loosely on Bartlett's training procedures, and was conducted during a one-week period in the fall of 1984, and again in the spring of 1985. Training occurred in five successive one-hour sessions during the students' regular CESL reading classes. The training covered four of Bonnie Meyer's (1975) major expository discourse types---the same four used by Bartlett (1978) and discussed in detail in Carrell (1984b).

Generally, Bonnie Meyer's research on expository texts has shown that there are a small number of basic types of expository organization (5 or 6) -- collection, description, causation, problem/solution, and comparison.

Collection and description may combine to be a collection of descriptions. These 5 or 6 types are not exhaustive or definitive, but Meyer's research has shown that there is good support for the belief that these are major text types and are significantly different types. Her research with native English readers, ranging from children to adults, has shown that when a reader possesses the appropriate formal schema for each of these text types, and when they activate that schema in interacting with a text, they remember far more content, both main ideas and supporting detail, than if they do not possess or do not activate the schema.

Basically the students were taught to recognize the top-level organization of a text according to Meyer's typology of text types. We taught them how to recognize the top-level organization of the text, and to use this information to guide their interaction with the text during reading, and to use this information to guide their recall of the text.

The sessions began simply, presuming no prior background and using several short and easy illustrative text passages; (sample training passages are published as part of the TESOL Quarterly article.) The sessions built during the week to longer and more subtle passages. All text passages were naturally-occurring texts, selected from a variety of sources, primarily social science textbooks and sources such as TIME magazine. Each session began and ended with reviews of the training program's objectives, and each session reviewed the previous session's main points.

The teaching style was intended to be highly motivating and engaging for the students and involved student interaction with the materials and individual corrective feedback. The teacher began by doing most of the talking, demonstrating, and so on, etc., but quickly shifted the responsibility for learning to the students.

The basic objectives of the teaching program were explicitly communicated to the students. Specifically, we explained to them that sometimes it does not matter how they read---for example, when they are reading for pleasure, but that at other times, it does. They were told that sometimes, especially as students studying English for academic purposes and headed for the university, they will be called on to read a lot of information and to remember it---for example, in preparing for exams and class assignments. We explained that the efficiency with which students can read under such circumstances was important; that if they can get the necessary information quickly and effectively, it is likely they will perform well and feel better about the task. We explained that over the training period, we would be teaching them a strategy for reading that should improve their understanding of what they read and their ability to recall it. We emphasized that by teaching them a little about the ways in which expository texts are typically organized at the top level, we hoped to teach them how to use this knowledge to improve their comprehension of what they read, as well as to teach them a strategy for using this knowledge to improve their recall of what they read.

During each training session, each student worked with a study packet, which was the focus of that session's activities. (Excerpts from the study packets are also published with the article.) Every day as the students left the session, they were asked to apply what they were learning to all of the reading they did until the next session. This was intended to get the students to use the strategy outside their ESL reading classroom, in other non teacher-supported reading situations.

The study packets included detailed explanations of the benefits of learning the strategy, along with checklists so students could monitor and regulate their own learning.

Control Sessions. While the experimental groups were going through the training sessions, control groups of similar students also received exposure to the same texts. However, they performed various linguistic operations with the texts, e.g., grammar exercises, sentence-combining, sentence analysis, work with discourse connectors, cohesion, and vocabulary work. They also focused on the content of the passages (for example, using the texts as a basis for question-answering and discussion), and they used the texts as the basis for various reading and writing assignments. The members of the control groups were told that the special texts they were being given, over and above their regular CESL Level 4 reading curriculum, were part of an instructional program designed to get them to think about (1) the linguistic aspects of various texts, (2) the linguistic operations they could perform on texts to become more aware of the connection between reading about a topic and writing about that same topic. The teaching style with the control groups also attempted to highly motivate and encourage these students as they worked with the texts.[1] The only thing the control group did not receive was the training on top-level rhetorical organization and the strategy for using that information as a basis for reading and recalling expository text. The control group did not receive training in any specific alternate strategy for use in the reading and recall of text. However, because Level 4 students at CESL are university-bound, considerable emphasis is given at that level to reading and writing for academic purposes, that is, reading to learn from text, using a variety of approaches. While the experiment might have been more tightly controlled if we had included a specific alternate training strategy for the control groups, we were more interested in comparing the top-level strategy training to what is commonly taught in an intensive ESL reading curriculum.

Testing Procedures. Both experimental and control groups were administered a pretest during the class period prior to training (Friday), and a post-test the class period following the week's training (Monday). A second post-test was administered to only the training groups 3 weeks after the first post-test, to determine the persistence of the training effect. Due to time constraints during testing, the pre-and post-tests covered only two of the four discourse types taught: comparison and collection of descriptions. Thus, the pre- and post-tests each included two texts, one of each of those two types. The length of the texts ranged from 230 to 281 words. To control for any effects of content schemata, all of the texts used in testing were about energy and environmental issues. The tests consisted of the subject reading each text, writing an immediate free recall, and identifying the text's overall organization by answering an open-ended question.

Scoring: Quantity of Idea Units Recalled. A priori each of the six texts used in the pre- and post-tests was analyzed into idea units. (An Appendix in the article illustrates the analysis of one of the test texts into idea units.) Basically, each idea unit consisted of a single clause (main or subordinate, including adverbial and relative clauses). Each infinitival construction, gerundive, nominalized verb phrase, and conjunct was also identified as a separate idea unit. In addition, optional and/or heavy prepositional phrases were also identified as separate idea units. Three

[1]The instructors of the two control and two experimental groups were these students' own regular CESL reading instructors, four different people. Thus, we were unable to control for any effects due to instructor differences. However, all four were highly regarded teachers with equivalent experience: two tenured teachers, one experimental, one control; two teaching assistants, one experimental, one control.

separate raters were used to arrive at the idea unit analysis of each text and all agreed on the final analysis.

Recall protocols were scored for the presence of each idea unit from the original text. In this scoring, each protocol was judged by two independent judges, with any discrepancies settled by a third judge. The pairs of judges achieved a reliability coefficient of r=.96. Because the number of idea units varied slightly from text to text (from 37 to 47), the number of idea units each subject recalled was transformed into a percentage of the total number of idea units in the original text.

Scoring: Quality of Idea Units Recalled. A priori each of the idea unit analyses of the six original test texts was organized into hierarchical levels (Meyer 1975, Meyer and Freedle 1984). Each idea unit was determined to be either a High-, Mid- or Low-Level idea unit, according to the following criteria:

High-Level: represents major ideas or main topics in the text.
Mid-Level: represents minor ideas or subtopics in the text.
Low-Level: represents minor detail in the text.

Organizing the idea units of each original text into a hierarchy enabled us to analyze the recall protocols in terms of the levels of idea units recalled and to determine whether the training was effective at all hierarchical levels or only for certain kinds of idea units---main topics, subtopics, or minor details.

Scoring: Organization Used. Each recall protocol was also analyzed to determine whether or not it utilized the text type of the original (i.e., collection of descriptions, or comparison) or not. To be classified as a collection of descriptions, the protocol had to have an overtly expressed topic plus associated comments on the topic; to be classified as a comparison, the overall structure had to contrast opposing points of views, either those of the original text or the subject's own point of view. The pairs of judges agreed 75 percent of the time in their scoring of rhetorical organization used.

Scoring: Organization Recognized. The open-ended questions were also scored on whether or not the reader had correctly identified the discourse type. The pairs of judges agreed 86 percent of the time in their scoring of rhetorical organization recognized.

Results

Although there may be some differences between the groups' performances on the collection of descriptions texts and the comparison texts which may warrant further analysis (see Carrell 1984b), the results reported herein are averaged across both text types.

First, the training enabled the experimental subjects to recognize the trained discourse types and to use them in their recall protocols. After training, the experimental group significantly increased in the proportions of those who recognized and used the text's top-level organization, whereas the control group did not. χ^2 tests of proportions for paired observations (Glass and Hopkins 1984: 291) yielded the following results:
Recognition: Experimental $\chi^2 = 10.71$, $p < .05$, Control $\chi^2 = .69$, n.s.;
Use: Experimental $\chi^2 = 10.29$, $p < .05$, Control $\chi^2 = 1.00$, n.s. Thus, the increases for the experimental group were statistically significant, those for the control group were not. Furthermore, the data from post-test 2 show the obvious persistence of the training effect, even three weeks after training. However, the basic point of the study was not just to

see if by so doing we could significantly facilitate their reading comprehension as measured by the amount of the original text they were able to recall.

A one-way analysis of variance prodedure shows no significant differences between the experimental group (Mean = 18) and the control group (Mean = 20) on the pretest ($F = 0.001$, n.s.). Thus, although we were testing in already existing classrooms and were not able to randomly assign subjects to experimental and control classes, there were no differences between these groups prior to training.

In a research design such as this, involving comparable pre- and post-tests, one might suspect performance on the post-tests to be somewhat predicable from performance on the pre-test. To the extent that this is the case, performance on the post-test cannot be attributed to the experimental activities and dictates use of an analysis of covariance procedure. Thus, the first thing to check is the correlation coefficient between pre- and post-test 1. An overall $r = .43$, $p < .01$, was obtained. Thus, the pre-test is significantly correlated with post-test 1, and an analysis of covariance is indicated. The second thing to check is whether there is homogeneity or regression. The test for homogeneity of regression was carried out using the GLM procedure in SAS. This procedure tested for the possibility of an interaction between groups and pre-tests. The interaction was not significant: $F = 1.14$, n.s. Because of the lack of a significant interaction, indicating homogeneity of regression, a One-Way Analysis of Covariance was carried out.

The results show that the mean on post-test 1 for the experimental group is statistically significantly larger than that of the control group, with the effects of the pre-tests controlled for. In other words, the training sessions with students in the experimental group were statistically significant in facilitating the total amount of information they could recall from the two types of texts they read. The persistence of the training effect for the experimental group is evident in post-test 2.

To determine whether the significant facilitating effect of the training applied equally to major topics (main ideas) as well as to subtopics and low level details, the results were analyzed according to the hierarchical levels of the idea units within the original text. Comparing results on pre-test and post-test 1 for both experimental and control groups, an analysis of covariance procedure for the High-Level idea units yielded an F-value, $F = 12.07$, $p = .0012$. Thus, the post-test 1 mean for the experimental group is statistically significantly larger than that of the control group for High-Level idea units. Results of the analysis of covariance procedure for Mid-Level idea units yielded an F-value, $F = 9.97$, $p = .0029$. Thus, the post-test 1 mean for the experimental group is statistically significantly larger than that of the control group for Mid-Level idea units. Results of an analysis of variance procedure for Low-Level idea units yielded an F-value as follows: $F = 6.48$, $p = .0146$. Thus, the post-test 1 mean for the experimental group is statistically significantly larger than that of the control group for Low-Level idea units.

These results show that the training sessions with the experimental group significantly improved their recall of High-, Mid-, and Low-Level idea units--main topics, subtopics, and supporting detail. Furthermore, the persistence of the training effect is evident in the scores obtained on the post-test 2, three weeks after training.

Discussion

The training experiment yielded promising results, demonstrating that explicit, overt teaching about the top-level rhetorical organization of

texts can facilitate ESL students' reading comprehension, as measured by quantity of information recalled. The results of the qualitative analysis show that the training facilitates recall of supporting detail as well as of major topics and subtopics. In addition, the persistence of the training was evident for as long as three weeks after training.

It should also be mentioned that student reaction to the training was extremely positive. Students expressed the view that they had learned a helpful technique which benefited them. The training students generally expressed more confidence in themselves as ESL readers.

The findings of this study are noteworthy, since, as was mentioned earlier, they not only support the earlier results obtained with an N of only 25 subjects, but they are even more robust.

Obviously, many more such training studies need to be conducted, (1) to refine the training techniques, (2) to determine the optimal length of training, (3) to discover whether there are differential effects at different proficiency levels of ESL, (4) to ascertain whether there are any differential effects of training due to differences in native language background, and (5) to determine the longer range persistence of the training effect. And, we might also add that we need to know more about the generalizability of the training effects to other reading behaviors.

As Carrell (1985) concluded, however, such training on discourse types is obviously only one part of a comprehensive instructional program in ESL reading comprehension. And as Tierney (1983:9) has said "It is easy to forget that the mastery of the strategy should not displace reading for meaning." A comprehensive instructional program in ESL reading comprehension should also include work in schema availability and schema activation (Carrell and Eisterhold 1983), meta-cognitive training (e.g., inference-awareness, analogy), comprehension monitoring skills, decoding skills, and so on [see Collins and Smith (1982) and Pearson and Gallagher (1983) for more on the latter]. Teaching the prototypical patterns of different texts would be inappropriate unless such instruction occurs in conjunction with helping students, in a number of ways, to acquire meaning from text.

REFERENCES

Bartlett, B.J. (1978). Top-level structure as an organizational strategy for recall of classroom text. Ph.D. dissertation, Arizona State University.

Carrell, P.L. (1984a). Evidence of a formal schema in second language comprehension. Language Learning, 34, 87-112.

Carrell, P.L. (1984b). The effects of rhetorical organization on ESL readers. TESOL Quarterly, 18, 441-469.

Carrell, P.L. (1985). Facilitating ESL reading by teaching text structure. TESOL Quarterly, 19, 727-752.

Carrell, P.L., & Eisterhold, J.C. (1983). Schema theory and ESL reading pedagogy. TESOL Quarterly, 17, 553-573.

Collins, A., & Smith, E.E. (1982). Teaching the process of reading comprehension. In D.K. Detterman & R.J. Sternberg (Eds.), How and how much can intelligence be increased (pp. 173-185). Norwood, NJ: Ablex.

Connor, U. (1984). Recall of text: Differences between first and second language readers. TESOL Quarterly, 18, 239-256.

Connor, U., & McCagg, P. (1983a). Cross-cultural differences and perceived quality in written paraphrases of English expository prose. Applied Linguistics, 4, 259-268.

Connor, U., & McCagg, P. (1983b). Text structure and ESL learners' reading comprehension. Unpublished manuscript.

Geva, E. (1983). Facilitating reading comprehension through flowcharting. Reading Research Quarterly, 18, 384-405.
Glass, G.V., & Hopkins, K.D. (1984). Statistical methods in education and psychology. (2nd edition). Englewood Cliffs, NJ: Prentice Hall.
Gordon, C.J. (1980). The effects of instruction in metacomprehension and inferencing on children's comprehension abilities. Ph.D. Dissertation, University of Minnesota.
Johnson, N.S., & Mandler, J.M. (1980). A tale of two structures: Underlying and surface forms in stories. Poetics, 9, 51-86.
Kintsch, W. (1974). The representation of meaning in memory. Hillsdale, NJ: Erlbaum.
Kintsch, W., & van Dijk, T.A. (1978). Toward a model of text comprehension and production. Psychological Review, 85, 363-394.
Mandler, J.M. (1978). A code in the node: The use of a storm schema in retrieval. Discourse Processes, 1, 14-35.
Mandler, J.M., &Johnson, N.S. (1977). Remembrance of things parsed: Story structure and recall. Cognitive Psychology, 9, 111-151.
Meyer, B. J.F. (1977a). What is remembered from prose: A function of passage structure. In R.O. Freedle (Ed.,), Discourse production and comprehension (pp. 307-336). Norwood, NJ: Ablex.
Meyer, B. J.F. (1977b). Organization of prose and memory: Research with application to reading comprehension. In P.D. Pearson (Ed.), Reading: Theory, research, and practice (pp. 214-220). Clemson, SC: National Reading Conference.
Meyer, B.J.F., Brandt, D.M. & Bluth, G.J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. Reading Research Quarterly, 16, 72-103.
Meyer, B. J.F., & Freedle, R.O. (1984). Effects of discourse type on recall. American Educational Research Journal, 21, 121-143.
Mosenthal, J.H. (1984). Instruction in the interpretation of a writer's argument: A training study. Ph.D. dissertation, University of Illinois at Urbana-Champaign.
Pearson, P.D., & Gallagher, M.C. (1983). The instruction of reading comprehension. Technical Report No. 297. Champaign, IL: University of Illinois, Center for the Study of Reading.
Reutzel, D.R. (1985). Story maps improve comprehension. The Reading Teacher, 38, 400-404.
Rumelhart, D.E. (1975). Notes on a schema for stories. In D. Bobrow & A. Collins (eds.), Representation and understanding: Studies in cognitive science (pp. 211-236). New York: Academic Press.
Rumelhart, D.E. (1977). Understanding and summarizing brief stories. In D. LaBerge & S. J. Samuels (Eds.), Basic processes in reading: Perception and comprehension (pp. 265-303). Hillsdale, NJ: Erlbaum.
Short, E.J. (1982). A self-instructional approach to remediating less skilled readers' use of story schema, causal attributions, and expectations for success. Ph.D. dissertation, University of Notre Dame.
Singer, H., & Donlan, D. (1982). Active comprehension: Problem-solving schema with question generation for comprehension of complex short stories. Reading Research Quarterly, 17, 166-186.
Taylor, B. M. & Beach, R.W. (1984). The effects of text structure instruction on middle grade students' comprehension and production of expository text. Reading Research Quarterly, 19, 134-146.
Thorndyke, P.W. (1977). Cognitive structures in comprehension and memory of narrative discourse. Cognitive Psychology, 9, 77-110.
Tierney, R.J. (1983). Learning from Text. Reading Education Report No. 37. Champaign, IL: University of Illinois, Center for the Study of Reading.

ISSUES IN DEFINING AND ASSESSING READING PROFICIENCY: THE ACTFL GUIDELINES

James F. Lee

University of Illinois
Urbana-Champaign, IL

In 1981, the American Council on the Teaching of Foreign Languages (ACTFL) published its Provisional Proficiency Guidelines (generic and language-specific descriptions) for the four language skills (speaking, listening, reading, and writing). In 1986, ACTFL published its updated and sometimes expanded Proficiency Guidelines (general descriptions only). The changes in the descriptions were by and large improvements. This paper focuses on the reading guidelines with specific reference to their theoretical underpinnings.

The 1981 descriptions of levels of reading proficiency were dominated by references to control over specific grammatical items which paralleled the traditional sequence of grammatical items found in most foreign language textbooks. Equating difficulties in reading with control over grammatical structures was typical for that time period yet did not hold up under the light of empirical investigation (Lee 1987). The reading process as described in the 1981 Guidelines was essentially a bottom-up process where meaning was accessed via control over the linguistic features of the target language. In 1986, however, the descriptions refer to top-down processes as well as bottom-up ones and references to specific grammatical items have been eliminated.

The present paper examines the 1986 Guidelines for reading proficiency along three lines for although they have been updated, certain problems remain particularly with respect to a definition of reading and to the approach to the assessment of reading. First, given the role assigned to background knowledge, reading is defined as a bottom-up process where top-down features compensate for underdeveloped linguistic competence rather than as an interactive process whereby meaning is found not only in the text but also created by the readers. Second, by assessing levels of reading proficiency via a hierarchy of skills, comprehension is implicitly defined as a quantifiable product rather than as a process. Third, by defining levels of reading proficiency according to text types without any reference to the readers themselves, the Guidelines account for only a partial model of reading.

ROLE OF BACKGROUND KNOWLEDGE

Implicity as well as explicitly the Guidelines state that background knowledge and topic familiarity compensate for underdeveloped second language linguistic proficiency. In particular, a Novice High reader "may be able to derive meaning from material at a slightly higher level where context and/or extralinguistic background knowledge are supportive." At the intermediate levels the reader brings to the texts "personal interest and/or knowledge" but it is only at the advanced level that "comprehension derives not only from situational and subject matter knowledge but from increasing control of the language." Even the advanced plus reader can compensate for a lack of linguistic development by being able to "follow essential points of written discourse at the Superior level in areas of special interest or knowledge." Almost contradictorily, it is only at the superior level that "reading ability is not dependent on subject matter knowledge." The role assigned to background knowledge in FL reading is quite clear; it is a fallback/feedback mechanism by which certain individuals, as James puts it, "may appear to know the foreign language better than they do" (1986, p. 105).

Research on first and second language reading has shown that background knowledge and knowledge about the subject matter do indeed play a role in comprehension. The usual choice of wording is that background knowledge and topic familiarity enhance comprehension. The perspective that should be taken is not how comprehension is enhanced due to background knowledge, but how poor it is in the absence of background knowledge. Background knowledge has also been shown to lead learners away from the main points of the text (Bernhardt 1986, Lee and Ballman 1987) and to influence a reader's interpretation of the author's intent (Johnson 1982). In this view, background knowledge is essential for comprehension; that is, comprehension does not or perhaps even can not take place without it. "No matter how well a student may know a language, he cannot read in that language with good comprehension if the subject of the text is one he knows absolutely nothing about and therefore can have no real interest in. Comprehension is also always directed and controlled by the needs and purposes of an individual and therefore crucially depends on an adequate amount of what is sometimes called 'background information'" (Dubin et al. 1985, p. 6).

McNeil (1984) defines reading as an interactive process between the reader and text by which meaning is both found and created. The process of reading comprehension involves "actively constructing meaning among the parts of the text and personal experience. The text itself is a blueprint for meaning." The role of the readers and their characteristics is underscored by the research conducted by Roller (1985). She found that readers and writers disagree on the important ideas expressed in a passage because each has their own perspective on importance. Specifically, a writer's perspective on what the important ideas are is a function of text structure whereas a reader's perspective is a function of the individual readers' background knowledge. Background knowledge is, therefore, far more to the reading process than merely a compensation for linguistic control over the target language. To relegate background knowledge to a compensatory role is to ignore the important role individual readers have in creating meaning from the blueprint-text.

PROCESS/PROJECT DISTINCTION

The dichotomy between process and product is evident in the Guidelines for reading proficiency which focus on the product of reading and not on the process. Process is used here to refer to the act of reading whereas product is used to refer to the quantifiable assessment of reading skills.

One third of the descriptors for levels of reading proficiency are based on demonstrating a select number of reading skills. There is, however, no mention in the descriptors of any of the processes involved in reading.

As seen in Figure 1, reading skills such as inferencing and extending hypotheses beyond the text, are a function of higher levels of proficiency. In applying the ACTFL Guidelines to adult second language reading, one immediately questions why skills are hierarchically arranged. Adult readers are able to use all the skills in the ACTFL hierarchy in native language reading because there is no question of an adult's cognitive maturity. We must assume then, that ACTFL's position is related to L2 linguistic development. Their position could be stated as follows: L2 readers are unable to engage in higher-order skills because of limited target language linguistic ability. In weighing whether foreign language reading is a language problem or a reading problem, Alderson (1984) concludes that "it appears to be both...we do not know [sic] this yet, and the question needs further refinement and intense investigation" (p. 24).

Research on native language reading comprehension has been criticized for taking a narrow view of comprehension assessment. Skills such as knowing word meaning, comprehending the literal sense, recognizing mood and author intent are but a small portion of what can be comprehended by the reader (Hewitt 1982 quoted in McNeil 1984). Moreover, skills assessment will always be biased in favor of the text rather than the reader. Skills-oriented tests are constructed from a writer's (text structure) perspective. As mentioned above, Roller demonstrated clear differences between readers' and writers' perspectives of important passage ideas. Similar findings are reported for FL reading by Bernhardt (1986) who showed that in recalling a passage, a certain subject placed great importance on a fairly minor aspect of the passage. This aspect of the passage, however, had been stressed in the subjects' previous instructional experience. Similarly, Lee and Ballman (1987) report that subjects' recalls of a passage on the political and social aspects of medieval feudalism were dominated by only the political aspects. From a text-structure perspective, the subjects demonstrated poor ability in extracting the main ideas (an ACTFL reading skill). By assessing skills, ACTFL focuses on the product of comprehension and not on the process. This focus is limited, limiting and inherently biased against the readers who take as much with them to the text as the text itself contains.

Figure 1. Hierarchy vs. Taxonomy

ACTFL Skills Hierarchy		Grabe's Skills Taxonomy
Level	Skill	
0/0+	recognize memorized elements	1. perception and automatic recognition skills
1	skim, scan	2. vocabulary
2	decode, classify	3. syntactic skills
3	infer, guess, hypothesize, interpret	4. cohesion
4	analyze, verify, extend hypotheses	5. coherence
5	all those used by educated native reader	6. author's stance
		7. application skills
		8. meta-skills

Developed from Canale et al. (1984) and Grabe (1986).

In contrast to the ACTFL hierarchy of reading skills is Grabe's taxonomy of reading skills (also presented in Figure 1). In positing a taxonomy of reading skills, Grabe states that L2 readers can use the skills necessary to a particular context if they are taught to do so. That the ACTFL Guidelines, which are experientially based, are a hierarchy of skills, may have more to do with a tradition of language teaching than with the process of reading. Traditionally, FL curricula present the "entire" grammar of the target language in the first year and assume this grammar instruction allows for the development of sufficient reading ability to introduce the national literature in the second year. If in the ACTFL experience, early stage learners can not make inferences or extend hypotheses, their experience may be traceable to a lack of reading instruction and not necessarily to a lack of FL reading proficiency. On the other hand, cultural readings are a part of many FL curricula. However, in a study examining the effects of assessment task on comprehension scores, Lee (1987) found that the scores on five different tasks varied significantly. The very traditional method of assessing comprehension, content-based questions written and answered in the target language, yielded rather low scores when compared to other assessment tasks (e.g. modified cloze passage and native language content questions). If experience shows low ability to demonstrate certain reading skills, then "perhaps foreign language teachers have been underestimating learners' comprehension because of the way they have been assessing it" (Lee 1987). The effect of assessment task was also demonstrated in an earlier study (Lee 1986) in which it was found that the quantity recalled from a passage was greater when the learners wrote their recalls in their native language than when they wrote them in the target language. Quantifying comprehension (as in these studies) yields a product-oriented assessment (How much was recalled? and How many propositions were identified?). While such product oriented tasks do allow us to infer what the processes of comprehension are, there are severe limitations and some extreme variation in the results they yield.

TEXT TYPES

James has developed a hierarchy of text types based on the degree to which the reader and writer share background information. His analysis has been adopted by ACTFL as a hierarchical classification of texts which define skill levels. Without detailing his explanations and criteria, the text types he identifies are: enumerative, orientational, instructive, evaluative, and projective. The hierarchy of difficulty and sample texts for each classification are given in Figure 2. As Bernhardt (1986, p. 25) points out, assessing reading proficiency is by definition reader oriented not text oriented. If you accept reading as an interactive process, you must accept that readers of all levels of proficiency can interact (to some degree no matter how limited) with all kinds of texts. Adopting an interactive perspective on ACTFL's hierarchy of text difficulty, the crucial issue is to discover not only what constitutes a difficult text, but how the reader interacts with such a text. What distinguishes a Novice reader from an intermediate one when they read an evaluative text? How is the Intermediate reader able to bring meaning to the text whereas the Novice reader cannot (if he indeed cannot)? Research also needs to be conducted which empirically examines the interaction between the ability to demonstrate a particular reading skill and the text type being read. Are certain skills more easily demonstrated on certain text types? This is indeed an issue for the ACTFL Guidelines since they hierarchize reading skills and test types.

Kaya-Carton and Carton are presently involved in a project to develop and validate a test of reading proficiency based on the Guidelines. In characterizing the reading process, they posit a full and partial model.

Figure 2. ACTFL Hierarchy of Text Difficulty

Level	Text Type	Sample Texts
0/0+	enumerative	numbers, names, street signs, money denominations, office/shop designations, addresses
1	orientational	travel and registration forms, plane and train schedules, TV/radio program schedules, menus, newspaper headlines, tables of contents, messages and memos
2	instructive	ads and labels, newspaper accounts, instructions and directions, short narratives and descriptions, factual reports, formulaic requests on forms, invitations, introductory and concluding paragraphs
3	evaluative	editorials, analyses, apologia, certain literary texts, biography with critical interpretation
4	projective	critiques of art or theater performances, literary texts, philosophical discourse, technical papers, argumentation

Developed from Canale et al. (1984), James (1986), Omaggio (1986), and Phillips and Dandonoli (1986).

The former takes into account the complete set of variables which contribute to the specific interaction between the reader and the text. The latter, on the other hand, takes into account only text characteristics. The authors state (p. 96-7) that "it may be perfectly legitimate to take an initial position that the reader characteristics are not of immediate concern, the objective being not to diagnose why the individual can or cannot read at a particular level, but to determine whether an individual can interact with a text which characterizes a particular proficiency level." It appears counter-productive to examine a partial model of reading after a full model has been posited. Rather than examining the partial model, be it text-based or reader-based, all effort should be made to examine the full model. By assessing different levels of reading proficiency based on the interaction between reader and text the full model of reading proficiency is examined rather than just a partial model thereby obtaining a more accurate description of a reader's level of proficiency.

CONCLUSION

What ought to characterize the proficiency level of the reader is how he interacts with different types of texts, and not the text itself. At the outset I stated that I would examine the ACTFL Guidelines along three lines. The three lines were: the role of background knowledge, assessing comprehension only as the product of reading, and defining proficiency levels as a function of text type. These lines are not altogether parallel

and therefore do intersect. They intersect at the point where it is necessary to define what reading is. Inherent in the Guidelines is that reading is essentially a bottom-up (text driven) process where meaning lies mainly in the text and is accessed through the reader's linguistic competence. How very different this is from defining reading as an interactive process where meaning is both found and created. While ACTFL looks to improve and update the Guidelines this paper offers several constructive suggestions. Because the ACTFL Guidelines are being proposed as national standards of measuring proficiency, they should be subjected to rigorous empirical evaluations.

REFERENCES

Alderson, J. C. 1984. Reading in a foreign language: a reading problem or a language problem? In Reading in a foreign language, ed. by J. C. Alderson and A. H. Urquhart, 1-27. New York: Longman.

Bernhardt, E. 1986. Proficient texts or proficient readers? ADFL Bulletin, 17. 25-28.

Canale, M., et al. 1984. The testing of reading and listening proficiency: a synthesis. Foreign Language Annals 17. 389-92.

Dubin, F., D. E. Eskey, and W. Grabe. 1986. Teaching second language reading for academic purposes. Reading, MA: Addison-Wesley.

Grabe, W. The transition from theory to practice in teaching reading. In (eds.), Teaching second language reading for academic purposes, ed. by F. Dubin, D. E. Eskey, and W. Grabe, 25-48. Reading, MA: Addison-Wesley.

Hewitt, G. 1982. A critique of research methods in the study of reading comprehension. British Educational Research Journal 8. 9-21.

James, C. 1986. Language proficiency levels and the typology of text types. In Defining and developing proficiency: guidelines, implementations and concepts, ed. by H. Byrnes and M. Canale, 97-106. Lincolnwood, IL: National Textbook Company.

Johnson, P. 1981. Effects on reading comprehension of language complexity and cultural background of a text. TESOL Quarterly 15. 169-81.

Kaya-Carton, E. and A. Carton. 1986. Multidimensionality of foreign language reading: preliminary considerations in assessments. Foreign Language Annals 19. 95-102.

Lee, J. F. 1986. On the use of the recall task to measure L2 reading comprehension. Studies in Second Language Acquisition 8. 201-11.

_____. 1987. The Spanish subjunctive: an information processing perspective. Modern Language Journal 71.

_____ and T. L. Ballman. FL learners' ability to recall and rate the important ideas of an expository text. In Foreign language learning: a research perspective, ed. by B. VanPatten, T. Dvorak, and J. F. Lee, 88-97. Cambridge, MA: Newbury House. (in press)

McNeil, J. D. 1984. Reading comprehension: new directions for classroom practice. Glenview, IL: Scott Foresman.

Omaggio, A. C. 1986. Teaching language in context. Boston, MA: Heinle and Heinle.

Phillips, J. K. and P. Dandonoli. Testing receptive skills in a proficiency mode. Paper presented at the Northeast Conference. April 1986.

Roller, C. M. 1985. The effects of reader- and text-based factors on writers' and readers' perceptions of the importance of information in expository prose. Reading Research Quarterly, 20. 437-57.

WHAT COMPOSITION THEORY OFFERS THE WRITING TEACHER

Patricia M. Dyer

Widener University
Chester, PA

Composition Theory, taken as a broad concept, has two meanings: first, all of the theories of composition from the classical view to recent process-oriented views, and second, current competing methods, techniques, and notions of composition pedagogy. In its first meaning, Composition Theory refers to the classical view (Aristotelian rhetoric), the neo-classical view (current-traditional), and more recent process-oriented views (Expressive, Tagmemic, Dramatistic, Cognitive, and Social). In its second meaning, Composition Theory, still in transition from the 1960's, refers to an expanding set of methods of composition instruction, specific functions of student writers and their teachers, classroom practices, and the supporting positions on how learning to write takes place.

Both meanings of Composition Theory inform the college composition teacher. The first is durable and expandable; the second provides the forum for exchange of information both on effective teaching techniques and on research of philosophical, psychological, and sociological views which may support particular pedagogies. The second meaning, current notions in Composition Theory, describes a confusion of partial and developing theories, suggestions for teaching, reports on student writing processes, and observations on the relationship of composition assignments to real-world writing.

Following Richards and Rodgers' (1982) division of method into approach, design and procedure, one way to order the situation is to consider today's process-oriented views of composition theory as methods of teaching and learning composition. When Richards and Rodgers' configuration is adapted to composition pedagogy, approach entails both a theory of the nature of composition and a theory of learning to compose. Design sets forth the relationship of these theories to instruction and specifies selection and organization of content, as well as roles of learners, teachers, and materials. Procedure refers to the techniques and practices which result from approach and design.

From this perspective, the approach of classical Composition Theory included the arts of invention, arrangement, and style, which form the basis of modern rhetoric, together with memory and delivery, which concerned speaking and were not carried forward to modern rhetoric (Young and Becker, 1981). Part of approach in classical theory was the notion that students learned to write from study of models of writing and

descriptions of organization and structure. The design of classical composition theory was built on invention through use of extrinsic and intrinsic arguments in developing a position. The procedure included delivery of written text as speech. Approach, design, and procedure can be characterized similarly for neo-classical (current-traditional) theory, an outgrowth of classical theory, which emphasizes the written text as an end in itself, not as text for oral presentation.

Application of the approach-design-procedure analysis to current process-oriented views of composition reveals methods of teaching and learning composition which are not as complete as those of classical and neo-classical theory. In these views the focus shifts from arrangement (structure and development of discourse) to invention (discovering the subject matter of discourse) according to Young (cited by Harrington et al., 1981). Five process-oriented views can be identified: expressive, tagmemic, dramatistic, cognitive, and social. The expressive view, known also as Pre-Writing, sees invention as a process of self-actualization of the writer. Introduced by D. Gordon Rohman and Albert O. Wlecke in 1964, Pre-Writing has been developed by Coles, Elbow, Stewart, Macrorie, and others (Faigley, 1986; Harrington et al., 1981). Another view, tagmemic invention, has its origins in Kenneth Pike's work in tagmemic grammar (Harrington et al., 1981). First presented by Richard Young and Alton Becker in the 1970's, it has been further developed by Flowers, Williams, and D'Angelo. Tagmemic invention identifies four stages of inquiry and three heuristics to deal with them. A third process-oriented view, dramatistic invention, devised by Kenneth Burke, uses act, scene, agent, agency, and purpose as its heuristic (Harrington et al., 1981). Bycker and Anderson, Winterowd, Berthoff, Gibson, and Coles have contributed to this view.

The cognitive view, developed by Britton, Emig, Flower and Hayes, Kroll, Lunsford, and others, relates invention to cognitive processes (Faigley, 1986). Finally, the social view, which concerns itself with the social rather than individual nature of invention, emphasizes collaborative learning and writing. Bizzell, Bartholomae, Bruffee, Heath, Faigley (1986), Cooper (1986), and Reither (1985), among others, support this view. The three most visable of these process-oriented views, expressive, cognitive, and social, warrant analysis in terms of Richards and Rodgers' three-level format.

For the expressive view, characterized as the pre-writing method, approach encompasses a view of composition in which thinking differs from and precedes writing (Faigley, 1986). With respect to design, the student writer is self-actualized through the discovery process. The teacher's role is to facilitate the student's composing process. Procedure includes numerous pre-writing techniques such as freewriting, journal writing, and brainstorming.

In the cognitive view of process, approach associates stages in learning to write with cognitive processes. Proponents maintain that an individual constructs reality through language (Faigley, 1986). Here, design includes analysis of protocols of both student writers and experienced writers to provide explanation and example for students. The procedure for cognitive process introduces development of heuristics by student writers. Textbooks supporting the use of this method show the stages of development of student writing from brainstorming through several drafts of an assignment, including the writer's additions and revisions, as well as the writer's comments on these changes.

For the social view of process, approach treats language from the perspective of a society, that is, it sees an individual writer as a

constituent of a culture (Faigley, 1986). Specifically, during the writing process, the writer is involved with interlocking systems of ideas, purposes, interpersonal interactions, cultural norms, and textual forms (Cooper, 1986). A design for the social view is not clearly specified. However, both Bruffee and Reither explain that writer and writing may not be removed from the writing situation itself (cited in Cooper, 1986). They show that writing in the professions includes review of text by colleagues as input to the final text. In this sense it is collaborative writing. Procedure, for the social view, is also unspecified, with the exception of use of written and oral peer input in the form of review and reaction to content. All in all, in its contemporary meaning, Composition Theory encompasses five partially specified process-oriented methods with their accompanying theoretical approaches and particular classroom practices.

Composition teachers, therefore, need to recognize both the historical meaning for Composition Theory and the contemporary meaning. While at present it may be inprecise to speak of a current single and complete Composition Theory, composition teachers sense a dynamic as Composition Theory redefines itself. Appropriately, they select from the literature those procedures most promising for their teaching situations, thus acquiring information on classroom results which may be the basis for future research and ultimately contribute to the developing theory.

Composition Theory, incorporating both meanings from here on, provides information on arrangement, style and invention to both composition teachers and business and technical writing teachers. In A Theory of Discourse, James Kinneavy (1971) identifies four varieties of discourse by their aims: expressive, which expresses thoughts, feelings, and beliefs; referential, which refers to reality outside the writer; literary, which focuses on itself as artifact; and persuasive, which is intended to provoke action or evoke emotion (cited in Harris, 1979). In teaching both composition and advanced composition, college writing teachers typically assign expressive, referential, and persuasive writing tasks. Next, students consider literary discourse in introductory literature courses, where the focus changes to writing about literature, a slightly different writing assignment. Later, in creative writing classes, students write their own poetry and short stories, thus learning about the artifact. Harris applies this classification to the teaching of technical writing, a type of expository writing. She sees technical writing as primarily referential discourse, and allows for a special category, persuasive-referential writing, to cover business letters and recommendation reports.

Kinneavy identifies three types of referential writing: exploratory, informative, and scientific. For Kinneavy, exploratory discourse asks a question: informative discourse answers an implicit question; scientific discourse ". . . offers proof of a hypothesis" (quoted in Harris, 1979, p. 629). In college composition courses, writing teachers assign readings, which introduce students to exploratory and informative discourse as preparation for writing assignments. In reading journal articles as preparation for their research papers, some students encounter scientific writing; other students will meet scientific writing later in their college careers when they read assignments for courses in other disciplines. Texts designed for composition courses provide students with reading selections and writing exercises which give students experience in reading and writing expressive, referential, literary, and persuasive writing.

Kinneavy's three types of referential writing may be related directly to technical writing assignments, according to Harris (1979). She reports structuring her technical writing course to begin with informative writing so that the modes of discourse may be introduced early on. This enables her to explain to students how the modes can be adapted for exploratory,

scientific and persuasive-referential writing. Texts on technical writing frequently treat modes of discourse in early chapters (Carosso, 1986; Lannon, 1985; Mills and Walter, 1986; Roundy, 1985; Warren, 1985). However, chapters with explanations of persuasive-referential writing, generally precede chapters treating informative, exploratory, and scientific writing.

Classroom Applications

Three elements of Composition Theory, a process approach to writing, an application of rhetorical strategies, and attention to audience, provide a workable foundation for instruction of college students in composition courses. The first of these, using a process approach, draws from both the expressive and cognitive views. Students in writing classes need to learn how to look at a writing assignment as problem solving. Steps in problem solving include assessing the communication problem, identifying a possible solution, writing a first draft, soliciting input, reviewing the draft in order to revise the original plan, editing, and, finally, rewriting the text. Certainly, product is paramount for most students in freshman composition classes. For this reason, these students may not be easily engaged by some of the classroom procedures of the expressive view of process, because they simply want to turn in the completed assignment. However, for the student writer, developing a personal writing process is essential. Thus writing teachers stress the value of investing in the process to achieve a successful product. The cognitive view, another process-oriented approach, is applicable in college writing classes since students can profit from learning how writers successfully proceed through writing tasks. In Student Writers at Work, for example, Sommers and McQuade (1986) show student essays in development from beginning stages to final draft by including successive drafts with the student writer's comments and editing symbols. Similarly, Lannon provides students with writers' accounts of their writing processes on the job; the accounts include descriptions of brainstorming, revising, and seeking advice from colleagues as part of the writing process.

Student use of the computer in writing assignments fits well with the cognitive view of process. In practice, when students use a word processing program on a personal computer for their writing tasks, they can readily see the stages they have worked through by reviewing hard copies of successive drafts. Students using Volkswriter Deluxe Plus and Word report they are motivated by being able to block and move key phrases, sentences, and sometimes entire paragraphs to better locations in the text. Moreover, there is no doubt that spelling checkers in word processing programs assist students in editing their work. Some spelling programs locate identical words which appear next to each other in the text and ask the student to verify this situation. Spelling programs, however, can not deal effectively with homonyms or words which do not occur in their dictionaries, for example, technical terms and proper nouns.

Several computer programs emphasize pre-writing and therefore seem to build on the notions of the expressive view. For some students, programs like Think Tank, HBJ Writer, and Mindwriter, which aid in the planning process, may be helpful. These programs allow students to do such things as brainstorm on the screen, answer questions on their topic generated by the computer, and develop an outline for their work.

Frequently, students opt to run a style analysis program, for example Rightwriter or Grammatik. Use of a style analysis program is consistent with the cognitive view's emphasis on the importance to writers of thinking through changes, in effect talking themselves through the changes. The precursor to style analysis software for personal computers, Writers' Workbench, runs on a mainframe computer and provides writers with infor-

mation on such things as length of sentence, number of sentences which begin the same way, use of passive voice, and use of uncommon and overworked words or phrases (Wresch, 1985). Style analyzers designed for personal computers provide students with similar information and in some cases suggest alternatives, allowing students to make changes on the screen. In reviewing the alternatives suggested, the student is in effect considering feedback, albeit computer-generated feedback. Often students who disagree with the computer's suggestion take the time to explain to their teacher why they decided not to change their text and ask for the teacher's opinion on this decision. More often students solicit input from other writers working at nearby computers. Such interaction is facilitated by a computer writing lab arrangement which places two or three chairs near each computer, an arrangement which proponents of the social view of process would surely espouse because it promotes collaboration.

The third process-oriented approach under consideration here, the social view of writing, emphasizes the role of social interaction in the process of writing. Social interaction during the planning of written communication is generally the situation in the workplace, where writers seek advice from colleagues at various stages in the writing process. In the writing classroom, students working on writing projects in teams can benefit from the exchange with their peers and begin to develop some strategies for questioning their own writing. In a recent presentation on developing critical thinking skills, Nelson suggested that students can develop critical thinking skills and learn to use a higher level of reasoning in considering texts if teachers develop appropriate group projects. For example, students may individually write on a topic and then bring their drafts to a peer discussion group. They should take notes on their peers' contributions and evaluative responses during the discussion and incorporate these ideas when revising their drafts. Finally, one of the most valuable aspects of use of a process approach in a writing class is helping students develop a self-monitor function, not simply for accuracy, but for tone, appropriateness, and overall assessment of effectiveness of problem solving. Students working together monitor their work by asking questions, providing suggestions, and offering evaluations to the members of the group, a process they learn to apply to their individual projects.

A second element of Composition Theory useful to writing teachers and their students comes from classical rhetoric: the application of rhetorical strategies to writing tasks. For the past two decades, many textbooks and readers developed for college composition classes have been organized according to strategies. More recently, however, such readers have included a thematic table of contents for teachers who prefer to organize their syllabus around themes. Conversely, many of the newer theme-centered texts also list their readings according to strategies, for teachers who like the selections but prefer to organize their syllabus around rhetorical strategies. This situation indicates that the shift to theme-centered texts, perhaps a response to research from the expressive view of process, has not replaced teachers' concern with teaching rhetorical strategies.

Some proponents of focusing on strategies suggest that introducing student writers to rhetorical strategies helps them identify choices available to them for developing papers. They liken these strategies to thinking techniques which students are accustomed to using (Tibbetts and Tibbetts, 1988). Others indicate that providing an overview of the modes of discourse (narration, description, exposition, and argumentation) and instruction in use of rhetorical strategies (enumeration, definition, process, comparison and contrast, analysis, cause and effect, and illustration) aids students in recognizing organizational patterns in written text (Arena, 1975). Students who become familiar with these patterns learn to use them in

their writing for freshman composition courses as well as for courses in other disciplines. For example, students reading a history exam learn to recognize essay questions which employ cause and effect or comparison and contrast and may be answered using these strategies (Carter and Skates, 1988).

For many college students, the writing class which follows the composition class is a course in business or technical writing. Students frequently arrive with the notion that the new formats (letters, memos, short reports), together with the new topics they will be using determine a form of short writing, used solely in business, which is completely unrelated to writing produced for other writing courses. However, it is essential that students see the relationship of business and technical writing to other types of writing. With some prompting, students may recall terms such as narration, description, exposition, and argument. Checking the table of contents in their text, they find familiar terms such as definition, process, classification, cause and effect, and comparison and contrast. By identifying the use of rhetorical strategies in their new writing course, students make the connection between their previous writing experiences and the new one.

The primary focus of technical writing courses is referential writing, typically found in short reports, feasibility studies, proposals, and longer analytical reports. In decoding the language of short problems and longer case studies, students can learn to identify strategies appropriate to organizing their response by posing questions: Will a proposed solution work for this purpose? Which of two solutions is better? Why does a certain problem occur? What are the effects of a particular action? Is a proposed solution practical? (Lannon, 1985). The answers students develop will suggest an approach to writing: analysis, comparison, causal analysis, reasons for--reasons against. Since solutions to complex problems covered in long reports generally require more than one form of analysis, students may properly apply several rhetorical strategies in formulating the written report, just as they did in writing the research paper for their composition class.

A third element of Composition Theory, attention to audience, also assists writing students. Audience, which is particularly important in both the cognitive and social process-oriented views of composition, is a prime consideration of effective writers. The writer may be writing for the self, the teacher, a broader audience, or an unknown, general audience (Tarvers, 1988). The most practical way for students to develop a sense of audience is to sketch briefly the characteristics of the audience they envision for each writing project.

For the student, developing a sense of audience is a challenging task. It requires the student to assume a role and to assign a role to a hypothetical reader or readers. The student needs to learn that the dynamic of the role relationship often includes both known and unknown tension between the two people in communication. Student writers need to try to imagine possible responses, negative as well as positive, to their writing. When the student writer determines a purpose for writing, including specific goals, he or she needs to consider what goals or agenda the reader might have. For example, if the composition class has been discussing the death penalty, the student essay writer may have peer readers who strongly oppose his or her position, as well as those who support that position. Students with the opposing view may read the essay with the intent to find loopholes in the argument. In another example, the business writer must consider whether the communication is solicited or unsolicited and has to focus on his or her relationship with both the first reader and second reader, keeping in mind the possibility of other readers.

Some composition teachers assign brief problem situations which require the student to envision the audience for a short essay or report. Here, the short problem, or case, provides information about the players and the problem, but requires the student writer to set up a hypothetical reader, based on his or her own input, before writing. Because group assignments facilitate analysis of audience during group meetings, such short cases are appropriately assigned to teams. Setting up teams composed of students from different disciplines contributes to the exchange (Tebeaux, 1985). To be most effective, these cases should provide a full rhetorical context, a purpose, an audience, and a role for the writer (Brockmann, 1984). For example, composition students may be considering the impact of proposed legislation which would permit replacement of a golf course with a shopping mall. One writer's purpose would be to inform the community of the benefits of such a change, while another writer might represent residents who oppose the change.

Case studies engage writing students in role play as a means of identifying goals and predicting strategies of both the writer and the reader. The instructor may direct the class discussion to identification and analysis of roles of participants in a controversy as well as to strategies available in conversational and written communication (Dyer, 1983, citing Di Pietro). Research by sociolinguists Hymes, Gumperz, Goffman, Sachs, Schegloff and others provides a departure point for discussion of role relationships, message interpretation, and verbal strategies (Dyer, 1983). In considering the social context of decision making, student writers begin to appreciate its bearing on selection of rhetorical strategies and choice of language for their written assignments.

Finally, there is much to explore in the contribution of the three elements of Composition Theory under consideration, a process approach to writing, application of rhetorical strategies, and attention to audience, with respect to what each offers writing teachers and their students. While these three elements are not the only major concerns of Composition Theory, they are three areas in which research clearly has made a valuable contribution to the efforts and direction of the writing teacher. Recently, information from proponents of the social view of process-oriented theory has directed researchers and teachers to the importance of considering the social context of writing. To this end, research on social interaction and discourse analysis from sociology and linguistics can enhance Composition Theory, particularly as it is applied to structuring collaborative writing assignments. Moreover, research on teaching methods which develop critical thinking by students can contribute both to Composition Theory and classroom instructional practices. The effects of applying research from these three areas to classroom writing instruction may indeed be improved instruction by teachers and greater facility with written language by students.

REFERENCES

Arena, L.A. (1975). Linguistics and Composition. Washington: Georgetown Univ. Press.

Brockmann, R.J. (1984). What is a Case? In R. J. Brockmann (Ed.), The Case Method in Technical Communication: Theory and Models (pp. 1-16). Association of Teachers of Technical Writing.

Carosso, R.B. (1986). Technical Communication. Belmont: Wadsworth.

Carter, B. & Skates, C. (1988). The Rinehart Handbook for Writers. New York: Holt, Rinehart and Winston.

Cooper, M.M. (1986). The Ecology of Writing. College English, 48, 364-75.

Dyer, P.M. (1983). Instructional Procedures for Implementing the Strategic Interaction Method in an Intensive English as a

Second Language Program. (Doctoral Dissertation, University of Delaware, 1982). Ann Arbor: UMI. 8309893.
Faigley, L. (1986). Competing Theories of Process: A Critique and a Proposal. College English, 48, 527-542.
Grammatik III. Computer software. IBM, 512K, PC-DOS or MS-DOS, disk.
Harrington, D., Keith, P., Kneupper, C., Tripp, J. & Woods, W. (1981). A critical Survey of Resources for Teaching Rhetorical Invention: A Review Essay. In G. Tate & P.J. Corbett (Eds), The Writing Teacher's Sourcebook (pp. 187-207). New York: Oxford Univ. Press.
Harris, E. (1979). Applications of James L. Kinneavy's Theory of Discourse to Technical Writing. College English, 40, 625-632.
HBJ Writer. Computer software. Harcourt Brace. IBM, 128K, PC-DOS 3.0, disk.
Kinneavy, J.L. (1971). A Theory of Discourse. Englewood Cliffs: Prentice-Hall.
Lannon, John M. (1985). Technical Writing. Boston: Little, Brown.
Mills, G. H. & Walter, J.A. (1986). Technical Writing. New York: Holt, Rinehart and Winston.
Mindwriter. Computer software. Fred Kemp, author.
Nelson, Craig E. (1989, June). Cognitive Development and Teaching: Facilitating Critical Thinking and Mature Valuing Across the Curriculum. Workshop, University of Chicago.
Reither, J.A. (1985). Writing and Knowing: Toward Redefining the Writing Process. College English, 47, 620-28.
Richards, J. & Rodgers, T. (1982). Method: Approach, Design, Procedure. TESOL Quarterly, 16, 153-168.
Rightwriter. Computer software. IBM, 128K, PC-DOS 3.0, disk.
Roundy, N.L. (1985). Strategies for Technical Communication. Boston: Little, Brown.
Sommers, N. & McQuade, D. (1986). Student Writers at Work. New York: St. Martin's Press.
Tarvers, J. K. (1988). Teaching Writing: Theories and Practices. Glenview, IL: Scott, Foresman and Co.
Tebeaux, Elizabeth. (1985). Redefining Professional Writing Courses to Meet the Communication Needs of Writers in Business and Industry. College Composition and Communication, 36, 419-28.
Thinktank. Computer software. Living Videotext, 1985, IBM, 128K, PC-DOS or MS-DOS 3.0, disk.
Tibbetts, C. & Tibbetts, A.M. (1988). Strategies: A Rhetoric and Reader. Glenview, IL: Scott, Foresman and Co.
Volkswriter Delux Plus. Computer software. Lifetree Software, 1983, IBM, 128K, PC-DOS 3.0, disk.
Warren, T.L. (1985). Technical Writing. Belmont: Wadsworth.
Word 5.0. Computer software. IBM, 512K, PC-DOS or MD-DOS, disk.
Wresch, W. (1985). Computer Uses in Prewriting and Invention. In S. Kleimann, E. Rice, & M. Scheltema (Eds.), Proceedings of the Third Maryland Composition Conference (pp. 137-151). College Park: Univ. of Maryland.
Writers' Workbench. Computer software. Bell Labs, VAX.
Young, Richard E., and Alton L. Becker. (1981). Toward a Modern Theory of Rhetoric: A Tagmemic Contribution. In G. Tate & P. J. Corbett (Eds.), The Writing Teacher's Sourcebook (pp. 129-149). New York: Oxford Univ. Press.

SYNTAX AND THE EVALUATION OF COLLEGE WRITING: A BLIND ALLEY

Carl Mills

University of Cincinnati
Cincinnati, OH

In 1977, Kellogg Hunt wrote that "in the mid-seventies ... the English teaching profession has a theory of syntactic development that covers a broad range of structures." He went on to say that "to the present time, the teaching of language has been guided almost exclusively by the rhetorician's intuition. But the theory of syntactic development reviewed here does not rest upon intuition alone; it rests on a solid body of experimental data" (Hunt 1977: 102).

Beginning with Hunt's pioneering work (1965), composition researchers and some applied linguists have increasingly maintained that readers respond, at least in large part, to syntactic complexity. Hunt determined that the t-unit, which he defined as "one main clause with all the subordinate clauses attached to it" (1965: 20), and which he later defined as "one main clause plus any subordinate clause or non-clausal structure that is attached to or embedded in it" (1970: 4), was the best measure of syntactic complexity that could be found. One can measure the average syntactic complexity of a text by calculating its mean words per t-unit.

In the writings of school children Hunt's contention that syntactic complexity increases with age has tended to be supported. With adult writers, even young adults, results have not been so clearcut. Steward (1978) found that surface structure clause length in words can be as important as t-unit length in the writing of adults. Still, Steward argued that t-unit length is the best indicator of syntactic maturity as writers move from high school to university. Working with sixth graders Higgs (1983) calculated a number of measures of syntactic complexity and found that narratives by high achievers tended to differ significantly in mean words per sentence from narratives by low achievers. On the other hand, Pendarvis (1983) found no differences in syntactic complexity between the writings of gifted and average seventh graders. O'Donnell et al. (1967) stated that for kindergarten and elementary school children "a simple, objective, and apparently valid measure of development of syntactic control is mean word-length of t-units." Quinn (1984), looking at "quality-rated writing of college pre-freshmen" found that there was no significant correlation between markers of "syntactic maturity" and judgments of writing quality. Similarly, Gebhard (1978) reported that "mean t-unit length, an established index of writing maturity, did not significantly distinguish the two quality-rated freshman groups. It would appear that

length of clause, reflecting skill in consolidating elements within larger structures, is a better indicator of quality than length of t-unit" (215).

While it is fairly clear that there is a correlation between increasing chronological age and increasing syntactic complexity in writing as long as we are talking about school children, there is no unequivocal evidence that syntactic complexity, whether measured by words per t-unit or words per clause, has any relevance to the evaluation of writing by college students.

Nevertheless, mean t-unit length has become a standard measure of writing quality even at the university level. Counting words per t-unit and using mean t-unit length to evaluate competing composition pedagogies has become a sort of cottage industry in some universities. In the 1970's the notion of syntactic maturity was combined with an already obsolete version of transformational generative grammar, the Syntactic Structures (Chomsky 1957) model, to develop a method of teaching composition, sentence combining, wherein students are expressly encouraged to combine many short main clauses, using normal English subordinating, that is, embedding, to create complex sentences. Morenberg, Daiker, and Kerek (1978) claimed that the technique of sentence combining at the college level led to writing that was both longer in mean t-unit length and judged independently better than writing of students who were taught by traditional methods.

Many rhetoricians, following Ong (1982), have argued, usually on the basis of shaky linguistic theory and precious little empirical research, that students from so-called "oral cultures" tend to produce syntactically simple written texts with widespread use of simple sentences and coordinate structures, while writers from literate cultures tend to produce texts with a great deal of subordination. But Halliday (1979) has argued that spoken discourse tends to be more complex syntactically than written discourse.

Beaman presents an impressive array of findings that show that "contrary to many previous assumptions, spoken narrative is on the whole just as complex as, if not more complex in some respects, than written narrative" (Beaman 1984: 78). Furthermore, her research supports Halliday's conjecture that "the types of complexities involved in the two modalities are different [in that] the increased lexical density of writing is balanced by a relatively greater number of subordinate clauses in speech" (Beaman 1984: 78).

Subjects of the present study

Student writings for the present study came from the Exit Examination of the Freshman English program in the McMicken College of Arts and Sciences at the University of Cincinnati.

The University of Cincinnati is a large urban state university enrolling around 35,000 students on 4 campuses. Most of the approximately 25,000 undergraduates on the main campus are required to take the Arts & Sciences Freshman English program. Freshman English is one of the few programs that applies across all undergraduate colleges of the university. In any given quarter approximately 2,700 students are enrolled in Freshman English courses.

In order to assure greater standardization in the literacy skills acquired by freshmen, the English department recommended in 1978-79 that all students be required to write an exit examination essay before they are certified as having completed the program.

At any given time, therefore, the Freshman English program's exams constitute a very large data base that can be used for research in composition, rhetoric, or applied linguistics. The present study is based upon the 177 exit examination essays written in December 1984.

A copy of the exit examination question for December of 1984 is given in Appendix A.

The essays are read and rated according to a holistic general impression grading system like that outlined in Cooper (1977). Multiple independent raters grade each exam on a four-point scale. Each essay is rated by at least 2 independent raters. In cases of disagreement that affect the student's mark on the exam, a third independent rater is assigned. Normally, about 30% of the essays require a third rater. For the December, 1984, exam a total of 71 (40%) required a third rater.

On the examination that served as the source of data for this study a total of 64 students (36%) failed the exam. Since a given exam can have either 2 or 3 raters, depending upon whether there is a disagreement among raters, I used the mean score of the two or three ratings for each exam. This is in line with actual exam grading practice, where the mean of the 2 or 3 ratings actually determines passing or failing on the exam itself. For the exit exam given in December of 1984, mean scores ranged from 1.00, total failure awarded by both raters, to 3.00, the second-highest mark possible under the grading standards outlined above. For all 177 exams the mean rating was 1.801 with a standard deviation of 0.562. More specifically, 29 essays (16.38%) received a mean score of 1.00; 35 essays (19.77%) received a mean score of 1.33 (a '2' and two '1's); 14 essays (7.91%) attained the barely passing score of 1.67 (two '2's and a '1'); 66 essays (37.29%) were judged passing if unspectacular at 2.00; one essay got a score of 2.33 (two '2's and a '3'); 21 essays (11.86%) scored 2.50 (a '3' and a '2'); and 11 essays (6.21%) achieved a score of 3.00

Copies of two sample examination essays are given in Appendix B.

Data Analysis and Results

In order to test various models of composition research, the 177 essays were analyzed from a number of perspectives.

First, it was decided to test the naive assumption that long essays are judged to be better than short ones. This is not exactly a trivial matter: Witte and Faigley (1981) found that low-rated essays tended to be shorter (average length of 270 words) than high-rated essays (average length of 647 words). On this particular examination essays ranged from 228 words to 1395 words, with the median at 571 words, the mean length at 608.497 words, and the standard deviation at 191.091.

There is a significant association between essay length and grade (chi-square = 29.46; $p < .01$). Examining cell frequencies shows, however, that it is not the case that long essays tend to receive high grades but rather that weak essays tend to be short essays, which accords with Witte and Faigley's findings.

Because the dominant paradigm in composition research is, following Hunt (1965; 1977), based on measurement of syntactic complexity, most of this research project was devoted to measuring syntactic complexity in various ways and attempting to correlate syntactic complexity with grade on these exam essays. Since the conditions under which the essays were produced were tightly controlled, since all writers had essentially the same training, the same preparation, and the same specification of the writing task, the exit examination seems to me to constitute an excellent means to test the efficacy of syntactic complexity as a measure of writing effectiveness.

Essays ranged in length from 12 sentences to 72 sentences, with the mean at 33.068. Mean sentence length, a measure that Higgs (1983) had found

useful for evaluating narrative writings of sixth graders, ranged from 12.49 to 34.84, with the mean at 18.82 and the standard deviation at 3.316.

The bivariate correlation between grade and mean length of sentence (MLS) was significant ($r = .227$; $p < .01$), but statistically significant does not necessarily mean the same thing as 'explanatorily important', and a correlation coefficient of .227 can hardly be described as spectacular. In fact, r-squared, the coefficient of determination is only 0.0517 for MLS, indicating that variation in MLS accounts for only a bit over 5% of variation in grade.

From Hunt (1965) forward, most researchers have argued that mean length of t-unit (MLTU), with t-unit described, once again, as a main clause with any associated subordinate (surface structure) clauses and any associated phrasal modifiers, is the best measure of syntactic complexity. In fact, many researchers argue that MLTU is the best objective measure of writing quality.

For this particular set of essays, MLTU ranged from 11.327 to 26.667 words per t-unit, with the average at 16.049. This is above the average MLTU (14.4) that Hunt (1965) had found for twelfth graders but lower than the figure (20.3 words per t-unit) for Hunt's "superior adults" (written in Harper's and Atlantic). Correlation between MLTU and grade was .278 ($p < .0001$). As it turns out, MLTU is the best single predictor of grade on these essays, but here again, it must be noted that a statistically significant correlation coefficient of .278 (r-squared = .077) is hardly spectacular.

Some researchers, notably Gebhard have suggested that for adult writers, especially university freshmen, "length of clause, reflecting skill in consolidating elements within larger structures, is a better indicator of quality than length of t-unit" (1978: 215). Here, mean length of clause (MLCL) ranges from 6.798, about equal to Hunt's fourth graders, to 16.118, higher than Hunt's "superior adults." Average clause length was 9.923 words. The correlation between MLCL and grade was .23999, ($p < .001$) slightly better than that for MLS and slightly lower than that for MLTU.

Table 1

Summary of Syntactic Complexity in Relation to Grade

Grade	N	Mean Words per Sentence	Mean Words per Clause	Mean Words per T-Unit
1.00	29	18.395	9.763	15.610
1.33	35	17.409	9.243	14.627
1.67	14	17.826	9.610	15.221
2.00	66	19.319	10.145	16.639
2.50	21	20.280	10.377	16.967
3.00	11	19.632	10.748	17.414

Other measures of syntactic complexity, number of clauses per sentence and number of clauses per t-unit, did not correlate significantly with grade.

When we attempt to evaluate the three measures of syntactic complexity as predictors of grade in a multiple regression analysis, we find that the multiple correlation coefficient is about what we might have expected (R=.2875) and R-squared, the coefficient of determination is .08266. And F=5.19659, which is significant ($p < .05$, df.=3, 173).

A stepwise multiple regression analysis shows that the best predictor of grade is MLTU, with MLCL a significant but distant second, and MLS not even reaching an F-level or tolerance level for it to be included in the regression. In fact, with the two best predictors in the equations (MLTU and MLCL), R=.2875, and R-squared is .08266.

Again, the correlation between even the best measures of syntactic complexity and grades given to writings by adult freshman on this particular exit examination, though significant, is not particularly strong.

Table 2

Intercorrelation Matrix: Grade and Syntactic Complexity Measures

	GRADE	MLS	MLTU	MLCL
GRADE	1.0000 p=****	.2274 p=.001	.2780 p=.000	.2400 p=.001
MLS		1.0000 p=****	.8509 p=.000	.4884 p=.000
MLTU			1.0000 p=****	.6672 p=.000
MLCL				1.0000 p=****

So far, following the usual practice in composition research, we have been looking at aggregate measures of syntactic complexity. Beaman (1984), in her examination of oral and written narratives drawn from The Pear Stories (Chafe 1980), has taken a more analytic look at coordination and subordination in spoken and written narrative discourse.

In the present study we might ask whether the relative abundance or scarcity of certain clause types has any effect on judgments of writing quality.

To answer this question I analyzed the syntax of the 177 exit examination essays, looking at some of the clause types that Beaman examined in The Pear Stories. I tabulated the amount of main clause coordination, whether joined by and, but, so, yet, or or, or "zero coordination", that is, main clauses joined by either a semicolon or a comma splice. I also looked at finite nominal clauses, chiefly clauses beginning with that, nonfinite subordinate nominal clauses (those beginning with to and those introduced by V-ing), adjectival clauses, that is relative clauses beginning with who, which, that, and 0, and adverbial clauses introduced by a number of subordinators such as when, because, etc.

Since absolute essay length, whether measured in words or in sentences, varies considerably, I calculated a frequency index by dividing the number of occurrences of a given sentence type by the number of sentences in the essay and multiplying the result by 1,000 to yield a measure of the rate of occurrence per thousand sentences of each clause type for each essay. I also calculated a frequency index that measured the rate of occurrence of each clause type per thousand words of text.

In all cases there proved to be no statistically significant correlation between rate of occurrence of these clause types and grade on the exit exam. For example, the number of coordinated main clauses ranged from zero to 18, but this had no discernible relation to grade as a measure of writing effectiveness. Similarly, the number of finite nominal clauses ranged from zero to 31 per essay, but this again had no discernible effect on grade.

Is it possible that some other measure might correlate better with expert readers' judgments of writing effectiveness? Perhaps we should look at the words that freshman writers use.

Besides the total number of words in an essay, I calculated the number of unique words, the actual number of separate lexical items, factoring out repetition, in an essay. Thus it was possible to calculate the amount of repetitiveness in each essay. The number of unique words, which ranged from 120 to 490 with the median at 265.333 and the mean at 270.039, was divided by the number of total words in the essay to produce a type-token ratio that could then be correlated with grade.

The type-token ratio, given as the percent of unique words to total words in an essay, ranged from 28.53% unique words to 63.504%, with the mean at 45.57%. Despite this wide variation in the repetition ratio, the correlation between repetition ratio and grade was nonsignificant ($r = -0.01344$; $p = .4295$). In other words, repetitiveness or variability in word choice had no significant effect on grade.

In Cooper and Odell's influential book Evaluating Writing (1977), the chapter immediately preceding Hunt's chapter is "Computer-Aided Description of Mature Word Choices in Writing," by Patrick J. Finn. Finn begins by restating Ellis Pages's view that we can gain an indirect insight into a writer's syntactic (and presumably, semantic and pragmatic) choices by examining the lexical items a writer uses. We do this by examining word frequency counts made from essays. For example, Page says, "we may be interested in the complexity of a student's sentences, in the branching or dependency structures which he [sic] has the maturity to employ.... But the sentence-parsing programs for computers which exist now are not completely satisfactory for our purposes. We might therefore hypothesize that the proportion of prepositions, or of subordinating conjunctions, constitute a prox [a variable that approximates] for such complexity" (Page 1966: 240).

Page also suggests that we might tell a great deal about a writer's style by calculating the "proportion of uncommon words" (1966: 242).

Page did his early work with high school students in Wisconsin who ranged from 13 to 17 in age. Interestingly, Page (1968) shows a correlation of 0.32 between essay length in words and grade.

For this study, I first decided to reverse Page's concern with uncommon words and look at the proportions of the most common words in the essays. Based on a stratified sample of 30 essays and choosing words that occurred at least five times in each essay, I found the most common ones to be: the indefinite article a or an, the definite article the, the conjunction and,

the negative not, the personal pronouns I and they, the pronouns-complementizers-expletives it and that, and the prepositions in, of, to, and for. Three verbs, do, be, and will, were also among the most common words. Student, a word virtually guaranteed by the topic choice and background readings, was the only noun.

The correlation between the proportion of student in an essay and grade was 0.2757, which barely missed attaining significance ($p = .07$).

The various forms of to be were correlated with grade fairly strongly ($-.325$; $p < .05$), apparently offering some support to teachers of composition who continually exhort their students to avoid passive and stative constructions with be.

None of the other common words correlated significantly with grade.

To test Page's contention that we can gain an indirect measure of a text's syntactic complexity by calculating the proportions of prepositions and subordinating conjunctions, I performed these calculations. Correlation between prepositions and subordinating conjunctions and grade was non-significant ($p = .07$) above, also did not correlate significantly with grade. The same holds true for the relative pronouns who and which ($p = .27$).

Since use of a great many coordinating conjunctions tends to reduce the average syntactic complexity of a text, one might predict, following Page's view, that and, but, nor, or, and yet should correlate negatively with grade. The correlation coefficient is -0.158, but this is far from significant ($p = .2$).

Finn (1977) proposes to extend Page's word frequency analysis by using, in addition to the proportions of certain words, the standard frequency index for each word as calculated in various corpus-based studies, chiefly Carroll, Davies, and Richman (1971). For example, "if a word has an SFI of 90, one would expect to find it once in every ten words (the word the has an SFI of 88.6). If a word has an SFI of 80, one would expect to find it once in every 100 words" (Finn 1977: 74). And so forth.

Finn's hypothesis is that unusual, but not freakish, word choices mark a "mature style." Based on this view, we might expect the mean SFI for a highly rated essay to be lower than for a lower rated essay. So, for the 30-essay sample here, I calculated the mean SFI for each essay. There was a negative correlation between mean SFI and grade, indicating that relatively uncommon word choices were valued. But the correlation was not significant or even close ($p = .276$).

Finn suggests that certain classes of content words be looked at in evaluating raters' responses to essays, though Finn (1977) reports no quantitative results. He suggests that researchers look at "Abstract Nouns, Verbs Denoting Cognitive Activity and Adjectives Judging an Abstract State" (Finn 1977: 85). There are some problems in applying quantitative analysis to Finn's classes because he gives only a few examples of each class. But I have tried to classify the words in these essays according to my interpretation of the examples Finn (1977) gives.

First off, despite Page's (1966, 1968) suggestion that we look at proportions of uncommon words, there was no significant correlation between grade and the proportions of any of the word classes listed by Finn (1977).

Looking at the SFI's of abstract nouns, abstract adjectives, and cognitive verbs, we find mixed results. The mean SFI for abstract nouns does not correlate significantly with grade in the 30-essay sample ($p = .19$), and

the correlation, which we could reasonably expect to be negative, since we are predicting that uncommon words would be highly valued, was positive (r = n.18).

For cognitive verbs the correlation was -.198, indicating that readers do tend to favor the use of relatively less common cognitive verbs, but here again, the correlation was not significant (p = .147).

For Finn's category abstract adjectives the correlation with grade almost attained significance (p = .0525), but here, as with abstract nouns, the correlation was positive (r = .3018), again indicating that common, rather than uncommon words were favored by expert readers.

As I mentioned above, there are problems in knowing just how to subject Page's and Finn's categories to quantitative analysis. But if my interpretation of their categories is correct, it is safe to say that the measures of lexical analysis that they suggest are not particularly good predictors of grade on this kind of writing task.

To conclude the discussion of lexical analysis, I must report some puzzling findings. The word frequency counts that I made for the 30 essays were produced by relatively inexpensive, relatively fast computer software. The frequency counts consisted of alphabetized word lists with frequencies given for each word in an essay. No contextual information was provided, but this seems to be in line with Page's and Finn's reports. This leads to a problem with the participial forms of cognitive verbs, particularly the ing-forms. For in English a verb ending in ing can function as a verb in the present progressive construction, as in The planes are flying, as a deverbal adjective, as in Flying planes [not grounded ones] can be dangerous, or as a deverbal noun, as in Flying planes [but not knitting sweaters] can be dangerous. But on a raw word list it is impossible to know the syntactic category of a given ing-form.

For this reason, I decided to exclude the ing-forms from my analysis of cognitive verbs. But since I had to identify them anyway in analyzing the word frequency lists, I created a catch-all category of ing-forms simply because I was forced to attend to them.

Strangely, the mean SFI of the ing-forms correlated with grade at .3819, the highest correlation attained in the lexical analysis (p = .019). As with the nearly significant correlation with abstract adjectives, the correlation is puzzling positive, indicating that readers seem to prefer relatively common cognitive verbs. But most puzzling is why the correlation of ing-forms with grade exists at all. Clearly, these lexical items require more study.

At this point it became obvious that there were some non-syntactic differences among the essays. Most notably, students writing the examination essays appear to have taken either of two interpretations of the statement of the writing task. Looking again at Appendix A, we can see that the "background" statement indicates that the students' former high school principals are proposing that all ninth through twelfth graders be required to attend a once-a-month, one-hour after-school seminar on dangers of substance abuse and coping with stress. But the actual statement of the "writing task" asks the students to write a letter to their former principals in which the student writers "agree or disagree with the argument that America's adolescents are seriously troubled."

Notice that the statement of the writing task does not require that the writers take a position on the proposed seminars. Read narrowly and literally, the task definition sets up a writing task that could be satis-

fied without arguing for or against the proposed seminars. But 154 (87%) of the student writers interpreted the writing task broadly to require that they take a position on the proposed seminars, with 121 (68.36%) arguing for the seminars and 33 (18.64%) arguing against them, while the remaining 23 students (13%) interpreted the task definition narrowly and argued the question of whether teenagers are seriously troubled without taking any position on the proposed seminars.

A one-way analysis of variance (ANOVA), using post-hoc pairwise comparisons evaluated under both the Student-Newman-Keuls and the more conservative Scheffe procedures, indicated that the position a student took on the seminars--for, against, or no position--had a significant effect on the student's grade. The mean score for all 177 essays was about 1.8. Students who argued explicitly against the seminars had a mean score of 1.8682; students arguing for the seminars had a mean of 1.8455. Needless to say, choice between these positions had no significant effect on grade. On the other hand, students who interpreted the task definition narrowly and took no discernible position on the seminars attained a mean score of 1.47, a failing score, which is significantly lower than the mean of the students who did take a position on the seminars ($p < .05$). Further, the highest score achieved in this last group of students (2.50) is lower than the highest scores (3.00) in the other two groups.

From this we can draw at least two conclusions. First, that writing teachers probably should advise students writing persuasive essays to take a clear, explicit, definite position on the issues raised in the exam question. Second, the framers of this particular examination question, in trying to be precise and helpful in their specification of the rhetorical context, unknowingly misled a substantial percentage of the students taking this particular exam.

A stepwise multiple regression analysis, with the variable of position entered in the equation along with the three best syntactic complexity measures yields a multiple correlation coefficient of $R = .36082$. More importantly, this analysis indicates that the position that a writer took on the issue of the seminars is the second-best predictor of the grade that the writer received on the essay. Only one measure of syntactic complexity, MLTU, turned out to be a better predictor than did position. That is, the pragmatics of language use, rules for getting things done with language, outweigh all but one syntactic variable as predictors of writing effectiveness.

Discussion and Conclusions

By now it should be obvious that, with regard to university freshmen writing persuasive essays under tightly controlled conditions, this body of experimental data is something less than "solid." Yes, significant correlations can be established between some measures of syntactic complexity, chiefly mean t-unit length, and assessments of writing quality. But a correlation coefficient of less than .3, however significant it may be statistically, does not indicate that syntactic complexity is a strong predictor of writing effectiveness. And the strong effect of including even one pragmatic variable, in this case position on the proposed seminars, indicates that the efficacy of surface syntactic complexity as a marker of writing quality is somewhat less than overwhelming.

Acknowledgments

This paper could not have been written without the work of students, faculty, and staff involved in the Freshman English Program at the University of Cincinnati. Special thanks should go to James Berlin, the Director of Freshman English, and Lucille Schultz, Acting Director. Helpful criticism

and suggestions have come from Barbara Risch, Brenda Johns, and Dr. Cathy J. Wheeler. Errors, omissions, and misinterpretations are entirely my own.

REFERENCES

Beaman, Karen. (1984). Coordination and subordination revisited: syntactic complexity in spoken and written narrative discourse. Coherence in spoken and written discourse, ed. by D. Tannen. Norwood, NJ: Ablex.

Carroll, John B., Davies, Peter, and Richman, Barry. 1971. The American Heritage word frequency book. New York: Houghton Mifflin.

Chafe, Wallace (Ed.). (1980). The pear stories: cognitive, cultural, and linguistic aspects of narrative production. Norwood, NJ: Ablex.

Chomsky, Noam. (1957). Syntactic structures. The Hague: Mouton.

Cooper, Charles R. (1977). Holistic evaluation of writing. Evaluating writing: describing, measuring, judging, ed. by Charles R. Cooper & Lee Odell. Urbana: NCTE. pp.3-31.

Finn, Patrick J. 1977. Computer-aided description of mature word choices. Evaluating writing: describing, measuring, judging, ed. by Charles R. Cooper & Lee Odell. Urbana: NCTE. pp. 69-89.

Gebhard, Ann O. (1978). Writing quality and syntax: a transformational analysis of three prose samples. Research in the Teaching of English 12, 211-231.

Halliday, M. A. K. (1979). Differences between spoken and written language: some implications for literacy teaching. Communication through reading: Proceedings of the Fourth Australian Reading Conference, ed. by G. Page, J. Elkins, & B. O'Connor. Adelaide: Australian Reading Assoc. II: 37-52.

Higgs, Rosalee Orma. (1983). The impact of discourse mode, syntactic complexity, and story grammar on the writings of sixth graders. DAI 45, 80-A.

Hunt, Kellogg W. (1965). Grammatical structures written at three grade levels. NCTE Research Report No. 3. Champaign: NCTE.

_____. (1970). Syntactic maturity in school children and adults. Monographs of the Society for Research in Child Development, no. 134, vol. 35, no. 3.

_____. (1977). Early blooming and late blooming syntactic structures. Evaluating writing: describing, measuring, judging, ed. by Charles R. Cooper & Lee Odell. Urbana: NCTE. pp.91-104.

Morenberg, Max, Daiker, Donald, & Kerek, Andrew. (1978). Sentence combining at the college level: an experimental study. Research in the Teaching of English 12, 245-256.

O'Donnell, Roy C., Griffin, William J. & Norris, Raymond C. (1967). Syntax of kindergarten and elementary school children: a transformational analysis. NCTE Research Report No. 8. Champaign: NCTE.

Ong, Walter J. (1982) Orality and literacy: the technologizing of the word. New York: Methuen.

Page, Ellis B. (1966). The imminence of grading essays by computer. Phi Delta Kappan 47, 238-243.

_____. (1968). The use of the computer in analyzing student essays. International review of education 14, 210-225.

Pendarvis, Edwina Dawn. (1983). A comparison of gifted students' written language with the written language of average students. DAI 45, 150-A.

Quinn, Dennis P. (1984). Rhetorical analysis of intellectual processes in student writing: linguistic cues in the quality-rated writing of college pre-freshmen. DAI 45, 2855-A.

Steward, Murray F. (1978). Syntactic maturity from high school to university: a first look. Research in the Teaching of English 12, 37-46.

Witte, Stephen P., & Faigley, Lester. (1981). Coherence, cohesion and writing quality. CCC 32, 189-204.

Appendix A

EXIT EXAM TOPIC

Background:

Your former high school principal has recently read an essay in U.S. News and World Report entitled "Troubled Teenagers." In that essay, the writer paints a bleak picture of American adolescents. He argues that they are growing up too fast, are unduly stressed, and are escaping "via drugs, suicide, crime, sex or just attempting to run away from it all." Concerned about the potential problems in her school, your principal is considering a new requirement for all students: once a month, all ninth through twelfth graders will attend a one-hour after-school seminar on the dangers of substance abuse and on positive ways to cope with stress. Some students she has talked with, however, argue that adolescents are not as troubled as the U.S. News essay suggests, and that such seminars would be a waste of students' time. To help her decide if the seminars are necessary, your principal is soliciting opinions from graduates of her school.

Your Writing Task:

Write a letter to your principal in which you agree or disagree with the argument that America's adolescents are seriously troubled. Use supporting evidence from the readings as well as from your own experience.

Appendix B

Sample Exit Examination Essays

[This essay received a grade of '1' (clear failure) from both readers.]

In observing our teenagers of today, I feel there are definitely serious problems among them which they fail to realize. There are many problems, such as drugs, suicide, and decrease in population of adolescents.

Drugs, in today's society, is with teenagers. Most teenagers don't realize the seriousness of this mind strickening problem in our society. The National Institute on Alcohol and Alcoholism asserts that, "about 15 percent of 10th to 12th graders and 11 percent of 7th to 9th graders are classified as problem drinkers."

A growing number of suicides by teenagers is another factor of life in which our teenagers should be in touch with. U.S. News and World Report indicates, "among 15 to 24 year-olds, suicide is the third leading cause of death after homicide and accidents." If teenagers were asked directly, have they ever felt lonely?, do they ever want to talk to someone about problems?, or do they feel fed-up with life?, most teenagers would say yes, not knowing that these are the major symptoms of suicide. Researchers of the National Institute of Mental Health reports that, "33 percent of adolescents have suicide impulses, up from about 10 percent in the 1960's."

Another index which high school students must realize is that the population of young people are declining. U.S. News and World Report says that, "population of Americans ages 21 and under has dropped from a high 80.8 million in 1970 to 74.9 million today." Statistics such as this should be presented to the teenagers to let them know that they are among this population which is declining.

Finally, if today's teenagers would slow down and realize some of the few factors that are slowly deteriorating our society, society in the future

will become a much better place to live. Teenagers are the hope of our future and their children.

[This essay received a grade of '3' (the second-highest possible) from both readers.]

In considering the youth problem, I can reflect back on my own days in middle and high school, and upon this reflection agree that many students are having difficulties in this area of their lives. These difficulties are the result of many different kinds of pressures exerted upon them. By far the most prevalent pressures are those from their peers, feeling they must conform or risk ridicule.

Unfortunately, in the last five years or so, one of these pressures has been to use drugs. Referring to a U.S. News and World Report study, seventy-two percent of high school seniors use alcohol, and thirty-four percent use marijuana. These are alarming figures, and from my personal experience, accurate, which more than demonstrates the need for an educational program.

The use of these drugs also helps students to cover up their problems. Among these problems are that of family life. Many times, difficulties at home prevent youths from building a solid foundation for their lives. A broken home can cause a great many problems for young people, including those of runaways, premarital sex, and violence. These only induce youths to drink more, take more drugs, and develop a vicious circle.

Many researchers believe that increased communication between parents and children is the most productive way to bring troubled teenagers back in line, with parents undertaking the effort. I believe this to be true, since parents are older, wiser, and more mature than teens who are not in full control of their lives.

In defense of those who believe our youth is in no need of help, studies by the National Center for Education Statistics do emphasize that many students know who they are and what they wish to do with their lives.

One major trend uncovered shows an increased emphasis on academics, and the importance of a well rounded education. They are beginning to realize that success, both personal and economic, depends on obtaining quality training and a solid, positive foundation for learning.

These attitudes can be derived from the general increase in grade levels, possibly due to improved study habits. Clearly, students are anticipating high levels of education in 1980 than they did in the early 1970's. Part of this change can be attributed to increased employment opportunities for those with higher education.

Concerning occupational expectations, several significant shifts occurred in seniors' values for the workplace. Among the higher rated values were important and interesting work, meeting friendly people,freedom to make decisions, and job security. Clearly, this study shows that many young people are prepared and anxious to begin their adult lives.

Consider a study by Lloyd Johnston and Jerald Backman of the University of Michigan's Institute for Social Research. They found that many youths are embracing traditional values. Defining traditional values would include a happy marriage and good family life, children, and a comfortable lifestyle. Religion and patriotism also seem to be making a comeback.

These indications are of a youth on the way up, but not without some problems at any given point in their lives. Consequently, a seminar to expose drugs for what they are is necessary. Parents would rather have their son or daughter learn about drugs in the classroom, than first hand out in the street. As a result, these youths will be able to recognize these peer pressure plays for their true value, worthless.

Appendix C

Grading Rubric for Exit Examination Essays

A '4' essay is well written. It demonstrates the writer's ability to use: (a) appropriate diction and syntax; (b) proper grammatical and mechanical usage for Standard Written English; (c) a coherent and effective structure; and (d) some awareness of the rhetorical relationship between the writer and reader. The essay develops an argument well. It (e) has a clearly defined thesis which addresses the exam topic; (f) has a thesis which evaluates different criteria for its subject; (g) presents serious and plausible evidence to support its thesis; and (h) is able to envision counter-arguments and offer evidence to refute them.

A '3' essay is not as well written as the essay described above. It may demonstrate or clearly suggest the writer's ability, but it may show one of the following problems: (a) occasionally inappropriate diction or syntax; or (b) an occasional but unsystematic grammatical or mechanical lapse; or (c) a coherent though less effective structure: or (d) a faulty awareness of the rhetorical relationship between writer and reader. Otherwise, this essay meets the requirements (e-h) above for the well developed argument of the best essays, though it may depend more on generalization than on detail to develop its argument and tend to propose "straw man" counter-arguments.

A '2' essay is only adequately written. It is pedestrian, thin in argument, acceptable but undistinguished in categories (a-d): diction and syntax, grammar and mechanics, structure, and rhetoric. This essay presents a cursory argument, superficially meeting the requirements (e-h) to address the exam topic, to evaluate different criteria, to present concrete evidence in support of its thesis, and to meet objections to its thesis. It is likely to depend upon generalization to develop its argument.

A '1' essay is inadequately written or unacceptably brief. It shows that the writer can address the topic, but it is terribly thin. It shows serious failures in either (a) diction and syntax, or (b) grammar and mechanics, or (c) structure, or (d) rhetoric. An especially weak essay may show more than one failure in (a-d). These essays do not adequately construct and develop an argument, failing to meet one (or more than one) of the requirements (e-h), to have a thesis, to evaluate any criteria, to marshall evidence to support the thesis, to present objections to the thesis or meet those objections on appropriate grounds. A '1' should also be assigned to essays only remotely related to the topic.

DISCOURSE ANALYSIS AND EMBEDDING DEPTH OF UTTERANCES:

CLAUSE ANALYSIS TECHNIQUE AS A MEASURE OF COMPLEXITY

Ghaida Salah

Al-Najah University
Nablus
West Bank, Jordan

INTRODUCTION

Clause analysis is a technique for measuring structural complexity. The significance of this measurement lies in its applicability to both written and spoken discourse. The clause analysis technique was first developed by W.A. Cook in 1969, and later expanded by Arena (1975).

The Structure of Speech Perception: The Clause as a Unit of Information

Bever hypothesized that the clause is the primary unit of information. In order to test his hypothesis, he conducted the so-called 'click' experiments in which he presented his subject with triple sentences which had similar surface structures but different deep structures (Bever as cited by Harman, 1974: 118-145). Bever's triple sentences are given below:

(a) The corrupt police can't bear criminals to confess very quickly.
(b) The corrupt police can't bear criminals' confessing very quickly.
(c) The corrupt police can't force criminals to confess very quickly.

In the experiment, the subjects were asked to wear headphones. The sentences were played into one ear and a click, which occurred during the word 'criminals', was played into the other. Subjects were asked to report where in the sentence they heard the click. In order to check that listeners did not utilize intonational or acoustic clues, the same recording of the segment 'criminals to confess very quickly' was attached to the three sentences, subjects tended to hear the click before the noun phrase, i.e., before the word 'criminals' where a Chomskyan deep structure suggests a structural break:

The corrupt police can't bear/criminals to confess very quickly.
The corrupt police can't bear/criminals' confessing very quickly.

In the third sentence, 'the corrupt police can't force criminals to confess very quickly', subjects were confused as to where to locate the click, because they could not decide where the structural break occurred. Their confusion was due to the fact that in deep structure, 'criminals' occurred twice with the structural break between the two occurrences:

The corrupt police can't force criminals/
criminals to confess very quickly.

This suggests, quite strongly, that people recover a deep structure when they decode a sentence. Other sounds that are unrelated to the sentence they hear will be ignored, and if these sounds have to be recalled, they will be reported as coming at the constituent boundaries.

From these experiments, it was determined that:

(1) Reaction time to clicks is faster at clause boundaries.
(2) 'Clicks' are accurately located in the language sample when they occur at clause boundaries.
(3) 'Clicks' are not located accurately when they occur at other than clause boundaries (Arena, 1982: 145)

Clause Analysis

This analysis is based on the tagmemic-grammatical protocol which posits the existence of the clause level between the sentence level and the phrase level in a five-level hierarchy. This hierarchy is: (1) Sentence, (2) Clause, (3) phrase, (4) word, (5) morpheme. Sentences are either simple, complex, or compound.

Steps in clause analysis:

The steps required in clause analysis are as follows:

(1) The corpus is segmented into single clauses, then listed on a line.
(2) Each clause is labled according to its basic type: (Mellon 1969)

iCl = intransitive clause
tCl = transitive clause
eCl = equational clause
pCl = passive clause
Min = minor type sentence
Code I Labels of Basic Clause Types.

An independent clause is marked with one of the labels of Code I. A minor-type sentence is an utterance that does not have any clause structure but has the function of a response to a question. The symbols used are:

A = a main clause
B = a subordinate clause embedded into A
C = a subordinate clause embedded into B
D = a subordinate clause embedded into C

(3) Mark all of the clauses listed according to type (code I)
(4) Independent clauses are labeled only according to type and not for any functional meaning, unless they are instances of directly or indirectly reported speech. In this case, they are marked according to their function. For example if the clause in question is a direct object, it is marked (DO).

Embedding Depth

The idea of clause depth is based on extended standard transformational theory. This theory states that a sentence, which consists of embedded clauses, is processed one clause at a time, starting with the lowest clause, followed by the next higher clause cycling upward until the main clause is reached.

In accordance with the cyclic rules, the processing time differs from one clause to another. Therefore, to calculate the depth of embeddings, numerical values are assigned to clauses in a corpus according to the following procedure: A = 1, B = 2, C = 3, and so on.

A clauses have a constant time in processing; B clauses take twice as long to process, etc. (Arena, 1982: 151).

To compute the embedding depth of any text, the following steps are followed: (See summary sheet)

(1) All clauses in a text are counted and the total number is recorded.
(2) All clauses are counted as to type, i.e., transitive, intransitive, equational, passive, or minor; furthermore, all clauses are marked as to whether they are main or dependent clauses.
The numbers are then recorded on the summary sheet.
(3) The frequency of each clause type is computed and recorded. The percentage of frequency of occurrence is derived by dividing the frequency of each clause type into the total number of clauses recorded in the analyzed text. (See summary sheet)

The use of embedding in discourse is closely related to the level of competence of the second-language learner, i.e., the more competence a learner enjoys, the higher his/her depth score is and the converse is also true. Simple and everyday topics do not call for complex clause structure compared to technical and more sophisticated topics which should reveal the creativity and competence of learners.

Subjects: Total number of students who participated in this study was: 62.32 juniors majoring in English in the Department of English at An-Najah National University, Nablus, West Bank and 30 First Semester Freshman students.

Setting: All the treatment was administered in the classroom during regular meetings in 1984-1985.

Task: All subjects were asked to divide into groups of two and develop scenarios on the following topic. Your father disagrees to your major at the university, how would you react. All dialogues were recorded, transcribed and then analyzed.

Summary Sheet Sample

No.	Symbol	Clause	Function	Type	Form
1	A	Hi, father			Min
2	A	Hi, son			Min
3	A	# I want(1)		tCl	
4	B	to know your opinion about my specializing in English#	DO	tCl	Inf
5	A	#Oh, my son,			Min
6	A+	I think		tCl	
7	B	English is not a good field for specialization nowadays.#	DO	eCl	
8	A	Why, father?			Min
9	A	# You can see		tCl	
10	B	Our country is full of teachers	Do	eCl	
11	B+	and we need money		tCl	
12	B+	and we are poor		tCl	
13	C	so it is better	R	eCl	
14	D	to search for another field#.		iCl	Inf
15	A	I am interested in English		eCl	
16	A	O.K.			Min
17	A	Thank you, father		tCl	

Totals

Embedding Depth: 1.59
No. of Clauses: 17

10 Main Clauses	4tCl	0iCl	1eCl	0pCl	5M
7 Dependent Clauses	2tCl	1iCl	4eCl	0pCl	0M

Primary Clause type employed — tCl — $\frac{6 \times 100}{17} = 35.3\%$

Secondary Clause type employed — Min — $\frac{5 \times 100}{17} = 29.4\%$

eCl — $\frac{5 \times 100}{17} = 29.4\%$

Tertiary frequent Clause type employed — iCl — $\frac{1 \times 100}{17} = 5.9\%$

Least frequent Clause type employed — pCl

(1) # Marks the beginning and terminal of sentences.

Embedding Depth

A clauses 10 x 1 = 10
B clauses 5 x 2 = 10
C clauses 1 x 3 = 3
D clauses 1 x 4 = 4

10 + 10 + 3 + 4 = 27 Value

$\text{Embedding Depth} \quad \frac{\text{Value}}{\text{No. of Cl}}$

27 ÷ 17 = 1.59

Table I. Embedding Depth of Dialogues

Freshman Students				Junior Students			
Group	No. of Cl	Value	Depth	Group	No. of Cl	Value	Depth
1	28	29	1.04	1	28	38	1.36
2	80	84	1.05	2	40	58	1.45
3	30	33	1.10	3	68	100	1.47
4	13	15	1.15	4	60	95	1.58
5	13	15	1.15	5	128	204	1.59
6	15	18	1.20	6	56	105	1.61
7	38	48	1.26	7	70	114	1.63
8	27	39	1.44	8	62	104	1.67
9	23	30	1.30	9	80	135	1.69
10	67	94	1.40	10	27	49	1.81
11	27	39	1.44	11	127	317	1.84
12	27	39	1.44	12	29	54	1.86
13	40	58	1.45	13	17	38	2.23
14	17	27	1.59	14	10	24	2.40
15	7	12	1.70	15	101	248	2.48
				16	30	86	2.86
Mean			1.30	Mean			1.84

Analysis of Data

When embedding depth was computed for the utterances produced by those subjects participating in this study, the following results were obtained.

If we consider the results of the lower level (Freshman students), we find that they became frustrated, shy and inhibited. Most of them hesitated then interacted in one-word utterances. Their performances were similar to those of children acquiring their first language. Clark and Clark (1977: 302) maintain that when children start to talk in their first language, they use one-word utterances. Most of their clauses were minor type such as: O.K., Yes, No, Why? Or equational such as: It's up to you. You are free, It is hard, etc. Or coordination which was predominant in their utterances. Hence, the mean of their embedding depth was 1.30 (see table I). These students scored low on embedding depth. I think that this result is attributable to their lack of competence.

We can say that silence prevailed, which means that the affective filter was high (Dulay and Burt, 1977). Krashen (1982) maintains that affective variables relate to success in second language acquisition. He adds that these variables can be placed into one of these three categories:

(1) Motivation. Most studies on motivation have reported that performers with high motivation generally do better in second language acquisition (Gardner & Lambert, 1972).
(2) Self-confidence. It has been proven that performers with self-confidence and a good self image tend to have better communicative competence. Canale and Swain's (1979) confidence is the ability to use language. They argue that a learner cannot have communicative confidence without communicative competence. In other words, they believe that competence is conducive to confidence. Lantolf's (forthcoming) notion of confidence is quite the reverse. He argues that learners cannot acquire communicative competence without having sufficient confidence in themselves and in their ability to interact. Recently, Savignon (1983: 45) expresses a similar interpretation of communicative confidence as that put forth by Lantolf. Corder (1981: 105) calls silence on the part of second language learners "risk avoidance." Corder believes that second language learners use the language when they are not inhibited by making errors and calls it "risk taking". But when learners are inhibited and avoid making errors, they remain silent.
(3) Anxiety: Low anxiety seems to be conducive to second language acquisition, whether it was personal or classroom anxiety.

While high anxiety is detrimental to learning, the proper amount of what Leontiev (1981) Calls operational tension may be necessary for acquisition to take place. Such tension provides learners with the motivation to confront the challenge of the situation.

When we compare the above results with those of Junior students, we can see the difference. This level showed more confidence and they were ready to participate. The mean of their embedding depth was 1.84 (see table I). In numerical terms, an average of 1.84 signifies that 84 percent more time is required to process the clauses of this level than is required for an average of 1.00, i.e., one clause per sentence without any embedding at all (Arena, 1982: 151).

A study by Mellon (1969) has stated that the accurate control and manipulation of sentence types is characteristic of syntactic fluency. This means that increased syntactic fluency indicates an increase in speaking or writing skills.

In the Mellon study, the T-unit was used to measure the students' development in their writing skills.

Arena (1975: 66) maintains that "the rules that Mellon employed in his segmentation of T-units were unclear as to how much of the student's writing they encompassed, and questionable as to why so many sentences of the corpus of students' compositions were discarded rather than included in the analysis of the student's writing. For example, of the seven rules for T-unit segmentation, four of the rules dealt with discarding:

(1) Each independent clause, including all constituent constructions, counts as one unit.
(2) Clauses of condition, concession, reason and purpose (although traditionally considered constituents of independent clauses) also count as separate T-units.

(3) Independent clauses occurring as directly quoted discourses count as T-units. Speaker tags are discarded.
(4) Orthographic sentence fragments count as part of the T-unit to which they belong.
(5) True sentence fragments resulting from the omission of a single word supplied. Other true fragments are discarded.
(6) Unintelligible word strings, vocatives, interjections, and various parenthetical or asyntactic expressions found in conversational writing are discarded.
(7) Independent clauses differing from preceding clauses only in their subject, and thus elliptical beyond their auxilliary, are discarded (Mellon 1969: 43).

The difference between the tool employed by Mellon, namely the T-unit, and the clause analysis of written English lies in that the latter does not discard any items and even accounts for minor sentences. Sentence type in this technique encompasses more than just the kind of sentences whether it is simple, complex, compound, or compound-complex. It also presents the type of clause and shows whether the clause type is transitive, intransitive, equational, or passive. Thus, in addition to an accurate clause analysis of the type used in this technique, a structural description of the kinds, the patterns, the variation of, and the balance among the various basic clause types which occur below the sentence level is presented and is available for computational purposes. This will provide us with a way to analyze and describe both syntactic fluency, balance and variation, in terms of the types of clauses that an author may employ throughout his writings. Hence, this technique enables us to specify an author's style whether it is expository, descriptive, narrative or technical. The following chart provides us with the frequency of occurrence of particular clause types in the above mentioned writing styles:

Style of Writing	Expected & Frequency of Occurrence of the 4 Clause Types*			
	Transitive	Intransitive	Equational	Passive
Expository	50%	20%	20%	10%
Descriptive	15%	30%	45%	10%
Narrative	20%	45%	25%	10%
Technical	10%	5%	20%	65%

*: ± 8% maximum variation permitted (Arena 1975: 88).

The above frequency of occurrence rates of the four basic clause-types for expository, descriptive, narrative, and technical writing are based on numerous clause analyses of published and unpublished examples of the above styles. Arena (1975) concedes that comparison of the figures contained in a clause analysis of an author's writing with the above frequency of occurrence rates will give an objective indication of the style of writing which the author employs, i.e., whether the writing is expository, descriptive, narrative, or technical, known as the author's "structural fingerprint", as well as whether or not there is balance or syntactic variation of the clause types as they occur in the writing. Arena adds, "The number of dependent clauses per complex or compound-complex sentence may also provide information pertaining to why a certain author may be more "difficult" or "easier" to read or to assimilate than others. To some, for example, Milton may be more difficult to assimilate because he employed 7.3 dependent clauses per such sentence in some sections of

Paradise Regained, while Ernest Hemingway may be viewed as "easier" to assimilate because he employed approximately two (2) clauses per complex or compound-complex sentence in The Old Man and the Sea.

Conclusions

(1) Clause analysis indicates the students' competence. Thus teachers can pinpoint the level of their students' errors. Tagmemicists believe that language is a system that includes a natural hierarchy of levels, e.g., morpheme, word, phrase, clause, sentence, paragraph, and discourse at large.

Examples for some of the types of composition errors that may be found at each level:

a. Morpheme-level errors: incorrect passive, tense, plural, and comparison; incorrect word form due to incorrect affixes, e.g., unprobable versus improbable.
b. Word-level errors: incorrect word form, diction, spelling.
c. Phrase-level errors: incorrect use of articles; adjective versus adverb distinctions, e.g., quick versus quickly. Misplaced adjectival modifiers.
d. Clause-level errors: sentence fragment; wrong embedded clauses, misplaced modifiers.
e. Sentence-level errors: faulty run-on sentences, (i.e., to continue without a break or a new paragraph), bad coordination of simple sentences.
f. Discourse level errors: weak organization, lack of smooth transition, faulty paragraph boundary, etc.

Errors decrease with increased proficiency:

(2) Clause analysis provides the teacher with precise evaluation of the students' progress. If one type, say, the passive clause type, is constantly missing, it means that the student is avoiding using it. Brown (1980) maintains that avoidance is a strategy that a learner of a foreign language resorts to in order to abandon a particular message he/she has started due to difficulty in expressing certain ideas in the target language. Some learners totally exclude some syntactic structures in their speech or writing. Avoidance is considered as one of contrastive analysis' drawbacks. Schachter (1974) has demonstrated that contrastive analysis is an inadequate tool for the study of a learner's errors because of the avoidance phenomenon inherent in error analysis. She adds that language learners will make errors on difficult items because they will simply avoid using the constructions altogether.

(3) Clause analysis can specify the accurate writer's idiolect or the writer's personal 'structural fingerprint' which can help in detecting instances of plagiarism.

BIBLIOGRAPHY

Arena, L. 1982. The language of corporate attorneys. In Di Pietro, 1982b. [Ed.] Linguistics and the professions: advances in discourse processes. Norwood, N.J.: Ablex Publishing Corporation. 143-154.

Arena, L. 1975. Linguistics and Composition. Washington, D.C.: Georgetown University Press.

Bever, T., Lackner, S., and Kirk, R. 1974. The underlying structures of

sentences. In Harman, G., 1974. On Noam Chomsky: critical essays. New York: Anchor Books. 118-145.
Brown, H.D. 1980. Principles of language learning and teaching. Englewood Cliffs: New Jersey, Prentice Hall.
Canale M., and Swain, M. 1979. Communication approaches to second language teaching and testing. Review and Evaluation Bulletin 1 (5).
Clark, H., and Clark, E. 1977. Psychology and language: an introduction to psycholinguistics. New York: Harcourt, Brace & Jovanovich.
Cook, Walter A. 1969. Introduction to tagmemic analysis. Trans-Atlantic series in linguistics. New York: Holt, Rinehart and Winston Inc.
Corder, S.P. 1981. Error analysis and interlanguage. Oxford: Oxford University Press.
Dulay, H. and Burt, M. 1977. Remark on creativity in language acquisition. In M. Burt, H. Dulay, and M. Finnochiaro (Eds.) Viewpoints on English as a Second Language. New York: Regents. pp. 95-126.
Gardner, R. and Lambert, W. 1972. Attitudes and motivation in second language learning. Rowley, Mass.: Newbury House.
Lantolf, J. (Forthcoming). Experiencing silent way at the university.
Mellon, John C. 1967. Transformational sentence-combining: A method for enhancing the development of syntactic fluency in English composition. United States Office of Education Cooperative Research Project No. 5-8418. Cambridge, Mass., Graduate School of Education, Harvard University.
Schachter, J. 1974. An Error in Error Analysis. Language Learning, 24, 2, 205-214.

THE ROLE OF LITERATURE IN TEACHING ESL:

STILL VIABLE IN THE 21st CENTURY

James R. DeJong

Delaware Technical & Community College
Terry Branch
Dover, DE

Despite much opinion to the contrary, literature, whether drama, narrative, or poetry, can be used effectively in the ESL classroom. Even in a world where increasing importance is placed on communicative competence, the use of literature in ESL has not been outdated. If one argued that it has become outdated, one would have to say that the use of dialogs (which dates from Plato and Aristotle's day) is also outdated. But the discriminating teacher knows that, in the great majority of cases, it is the use to which the material is put and the way in which it is dealt with that makes it either useful, or boring and outdated. Just because goals and objectives change then, does not mean that we should throw literature or any tool of the trade out, even if we think of them as so much extra baggage left over from the old-fashioned Grammar-Translation classes in which Homer's Iliad was used in translation exercises. Obviously, teachers today stress word for word translation less, and yet the possibilities for culture learning and empathy building through exposure to literature abound for those teachers who are willing to use the same well tried tools for a different purpose.

With good reason, language learning goals have swung increasingly toward the effective teaching of oral communication skills in the ESL and Foreign Language classroom. Our world is much smaller than it was twenty years ago due to increased jet travel and the wide variety of telecommunications available. The need to communicate, orally and in writing, whether in business, foreign diplomacy, or for tourist pleasure, with speakers of foreign languages in foreign cultures is now at an all time high the world over. But simply because language learning goals and needs have changed, does not mean that a tool such as literature is no longer a viable teaching aid. Our purpose is always to achieve the objective, -- in today's world, perhaps oral communication -- but we do not need to be married to old methods in order to use an old tool in the pursuit of the new objective. Literature can be adapted for use in the new methods coming out. In fact, it is my contention that despite literature's somewhat limited role in Grammar-Translation methods in the past, it can be and has been used effectively in today's methods emphasizing communicative competence, and indeed it is being used in many unique and interesting ways to obtain one of the specific language objectives needed in communicative competence, namely socio-cultural awareness and sensitivity, which I will call socio-cultural competence.

HISTORICAL BEGINNINGS

Let's start with a little history. Albert Marckwardt (1978) in his book on the use of literature in ESL comments that until quite recently literature was only used as a capstone for the foreign language learning experience. Examples he cites are the reading of FAUST, DON QUIXOTE, and works by Shakespeare for German, Spanish, and English classes respectively. The use of literature in this way, he sees as a clear hold-over from the time period of the Renaissance through to the nineteenth-century, when the Greek and Latin classics were studied not as media of communication, but simply for reading knowledge of a cultural masterpiece. In the same way modern language classics are often treated as mere objects of study to gain knowledge about one culture's masterpieces.

In the mid-nineteenth century, as the use of the vernacular in formal studies began to take precedence over the great Greek and Latin works in the curriculum, dispute arose over the place and purpose of works in German, Spanish, French, and English in the classroom setting. In countries where English is needed more as a language for study, (In Thailand and Malasia for example doctors and lawyers, etc. need to read English more than speak it) many teachers still use literature widely. But in countries where oral communication is stressed, many have dispensed with literature almost entirely, as being irrelevant to their new objectives. It is the non-use of literature in communicative approaches that I wish to deal with in this paper.

CONFLICTING VIEWS

The extreme view is a total relegation of literature to extra-curricular activity, such as Blatchford's (1972) statement, "the study of English literature is a luxury that cannot be indulged during the limited amount of time allocated to English" (p. 4 in Marckwardt) Marckwardt does state that Blatchford is dealing more with elementary and secondary schools, and that he allows for some use of literature where English is a second language (India, Nigeria, etc.), but still this is a rather extreme statement which I as a foreign language learner and teacher wish to dispute.

Blatchford's three main arguments are that classroom emphasis must be on functional use of the language, that teachers generally lack training in teaching literature (anyone can learn though), and that literature does not fit into the sociolinguistic and sociocultural understanding objectives common today. Blatchford states that teacher training in America is placing "more emphasis upon an understanding of the differences in register and sociolinguistic and sociocultural understanding that a teacher must have, rather than upon literature which does not contribute to the student's ability to function in society". (p. 5 in Marckwardt) Precisely on these three points though, I would argue that literature can be one of the most effective tools for gaining sociolinguistic and sociocultural understanding, that it can also serve an important purpose in teaching functional English, and that it is not hard for teachers with the help of a few books and articles to learn how to use literature effectively in their classrooms.

Blatchford and those who hold opinions similar to his, may think of literature study as the old fashioned boring translation and reading of stodgy texts under archaic teachers, as his stated opinion seems to indicate. Admittedly, the European and especially the German style of literary criticism with its attention to form more than meaning and content, which some teachers have inherited in the U.S., may also fall short of sociolinguistic objectives in learning. If all the attention is placed on the artificial forms of literature then obviously the student will learn only this aspect. (Even here though, after analyzing form, content and meaning can be dealt with.)

But what about the growing number of teachers who use the method of New Criticism, which emphasizes the content, meaning, and message of the work? What about opening the term literature up to include more than the Belle lettres? Is their anything wrong with reading a John Wayne style western novel, being engrossed in the action, and then discussing the American ideal of rugged individualism, or the strong silent type? Or what about choosing any of the wide number of modern works in drama, narrative, or poetry which involve dialog, and a very up to date use of everyday language, colloquialisms, phrases, proverbs, etc? Journalism is perhaps more practical in one sense, but that should not preempt literature from use. Not all literature is artificial as some have claimed. A close friend of mine from Spain who is studying at the University of Michigan, was reading Salinger's "The Catcher in the Rye" and commented on how it was helping him learn and practice some slang sayings which are still in common use today.

Regarding what I call cultural competence, for which Blatchford says literature cannot be used, I know of a surety that literature provides a unique window into the private family lives of hundreds of "native speakers", exposing their dreams, woes, hopes, and accomplishments, along with their everyday habits, likes, dislikes, and values. What more could a reader desire than to see in action all of the cultural traits he has heard of, and be able to stop, ponder and reread all or parts of an exchange between two or more native speakers. Really, after actually going to the TL country (often a very expensive business), and even after arriving and observing the culture first hand, reading, discussing, and writing about literature is one of the most effective means of actual culture learning that I have ever experienced. These and other points are what I will deal with in the rest of this paper.

COMMUNICATIVE COMPETENCE/PERFORMANCE

In the modern communicative approaches, all four language skills are emphasized, but speaking and listening take precedence over reading and writing. No teacher would go so far as to suggest that these two groupings are totally independent and unrelated to each other, for obviously the more contact a learner has with the language in any form, the faster he will progress. Often the main emphasis in speaking is placed on the communication of ideas by whatever means possible (even if mistakes are made), which is a change from the audio-lingual demand for absolute accuracy and purity of accent at every turn. The audio-lingual class is seen by some as a very sterile way of dealing with language, which aids little at times in actual conversational skills and communication. In the newer methods, listening skills are developed for the gathering of meaning, which seems to be a slight change in intent from the taped language labs where word for word understanding was often the required performance.

As for reading and writing, new ways of dealing with them have also developed. McKay (1982, p. 531) takes her ideas about reading from Widdowson, who thinks of reading "not as a reaction to a text, but interaction between the writer and the reader mediated through the text." This is quite a change from the past, for it regards reading not as passive and receptive only, but as active participation between individuals, which sounds much more like an ideal of the communicative approaches. Writing too, has taken on a different light. Zamel (1982) mentions that "writing is a process through which students can explore and discover their thoughts and ideas,..." She continues by saying that the process of writing, reading what is written, and then rethinking and rewriting, is like an inner dialog. This intense contact with the language is what learners often need, in order to sort through and practice using what they have learned or acquired. She bases her article on writing themes, using literature as a springboard for ideas and reactions. Inner dialog, and

what could be called deep thinking, seem to be what methods such as Suggestopedia use to a great effect for vocabulary retention and long-term memory. (I personally have found that when I have had to think and work for long hours with a foreign language in writing themes, my oral communication has indeed improved as a crossover from the reading and thinking. Specific vocabulary items and a feel for how and when to use them, has been a direct help in my communicating in the TL. McConochie [1985] mentions a similar experience with some of her foreign students.)

One of the major new components of the idea of communicative competence, is the sociocultural one. Now more than ever teachers and students are realizing that all the linguistic accuracy in the world just doesn't get the job done if you smell wrong, make the wrong kind of eye contact, stand too close or too far away, pause too long between phrases, etc. This has probably developed out of the dire experiences in the TL country, which so many more people have had now that world travel has increased. The fact that communication can break down despite perfect linguistic form and accuracy, has turned the sociocultural aspects of language into one of the new language learning objectives, as Blatchford's quotation pointed out.

LITERATURE IN THE CLASSROOM

Regarding the use of literature in ESL teaching, several authors have expressed themselves in favor of it and I think it can be adapted easily to meet the new objectives. Arthur (1968) in McConochie (1985, p. 206) reminds us that, " people tend to remember best what interests them deeply, rather than what they are told to to remember or are consciously trying to remember." McConochie also states that in her dealings with narrative poetry in class, she has found that "the language of literature,... stimulates readers to create meaning - to make sense of a text - by calling on their own experience just as people do in learning and using language naturally." Collie (1987, p.2) says that "literature, which speaks to the heart as much as to the mind, provides material with some emotional colour, that can make fuller contact with the learner's own life, and can thus counterbalance the more fragmented effect of many collections of texts used in the classroom." McKay (1982) says that "for some students, literature may provide the affective, attitudinal, and experiential factors which will motivate them to read." Obviously, reading is not the primary goal of communicative competence, but these statements show how literature can be used to provide the interest, motivation, and contact with the TL needed for faster learning, which is one of the prime objectives in the new communicative approaches.

I am sure that for decades, and even centuries literature classrooms in Europe and the U.S. were the epitome of boredom. This seems to explain the frequently negative reactions of many today when the use of literature in ESL is mentioned. In answer to this problem, McKay (1982) says that she found several factors which make the use of literature enjoyable to her students. She picks stylistically uncomplicated pieces that are not too difficult for the level of the ESL students she is teaching, even if this means using young adult literature.* She always approaches the piece first for aesthetic appeal, and meaning. And finally she starts with the conflict and begins discussion from there. (She does not place heavy emphasis on literary forms, figures of speech, etc., if it does

---

*(Something I have not yet tried but have thought a lot about is using children's literature such as Dr. Seuss, Curious George, and Little Red Ridinghood for beginning elementary-aged ESL students. This would give students a taste of the literature that the target language kids have read, and would be enjoyable as well, as a vehicle for introducing cultural contrasts.)

relate directly to meaning). Both McConochie (1985) and Spack (1985) have a systematic approach to actually studying the pieces. In their articles they outline how to preview the work and author, then to read, and then through in-class discussions, begin to get at the meaning. The interaction in class, as well as writing themes, give ample opportunity to think, discuss, and write in the target language all of which aid in communicative competence. This systematic approach reduces student anxiety, and allows for teacher guidance and input on cultural matters as well as linguistic matters, all the while encouraging the student to strive for comprehension of the many new linguistic items and forms before him. Marckwardt also comments on the necessity of bringing the students to a point where they will be moved by what they read (which he admits does not happen every week) and that they be granted opportunity to react to the work and apply it to life.

The consensus of opinion among many of those who use literature in the classroom, seems to be then, that given the proper selection of literary works, for length, difficulty, student interest, and applicability to the given objective; and a classroom approach to the literature based on emotional and aesthetic appeal, students have not had excessive trouble becoming interested in the work and really feeling that it applies to their life, or can be related to their own life experiences. This personal interaction and contact on a deep level with certain literary characters or ideas is obviously a help in learning how to use a language in communication. From my own experience and that of these authors, I feel it is safe to say that literature is a prime tool for motivating meaningful learning, and interaction on several levels.

BENEFITS OF LITERATURE FOR COMMUNICATIVE COMPETENCE

STRUCTURAL LEVEL

At the structural level, no amount of reading literature will substitute for the actual study of grammar, the actual use of the language in strategic interaction, or whatever method is used in other communication based approaches. A certain level of fluency and language competence can be gained during the first two years by these methods, but short of going to the target language country, how can a learner get the necessary practice in sociocultural variables, register, space, and actual conversational language use? I have always felt, ever since I began studying German and Spanish, that literature was my first chance to actually see the language in use among the natives. Granted I wasn't in Berlin or Madrid, but I was reading something written by Germans for Germans about Germans in Germany, and in German. The same for Spanish. This, over the years, has been one of the greatest sources of structure review and practice that I have had aside from living in both countries and actually interacting with native speakers. It could be argued that this has meaning only for one who has learned by the Grammar-Translation method, and therefore needs constant review and mental reassurance for the myriad individual structures floating around in his brain. But the benefits I have seen were not merely in structure review. Rather they were in language use and context appropriateness for the language that I already knew.

McKay (1982) supports this idea in her article on Literature in the ESL Classroom. She states that:

> literature presents language in discourse in which the parameters of the setting and role relationship are defined. Language that illustrates a particular register or dialect is embedded within a social context, and thus, there is a basis for determining why

a particular form is used. As such, literature is ideal for developing an awareness of language use.

(Page 530)

Precisely at this point of awareness of register and dialect, many classroom lessons fail, as can be attested to by the many exchange students in Germany who have addressed their University Professors with the informal DU and have then wondered at the look of consternation on their fellow classmates' faces, not to mention the dark clouds rising on the faces of professors who dislike or just barely tolerate Americans. As for dialect, during my stay in Vienna, Austria as a student, one of my prime sources for learning the local Viennese street slang, which I needed for survival and enjoyment, were the plays and novels written by local Viennese authors which contained dialogs written in authentic dialect and slang. I saw the words, practiced the words, understood the words and then began to use the words (for fun usually) in appropriate situations. These kinds of things can be dealt with very effectively through literature.

VOCABULARY

On the question of vocabulary, there are opposing views. Some say that the quantity and difficulty of new vocabulary items in literature is a formidable and insurmountable task for language learners. Marckwardt (1978, p. 69) says it is "important that the student be able to read smoothly and easily without constant thumbing of a glossary or dictionary." His preference for waiting for the fourth and fifth years of language instruction, is sort of a middle of the road acceptance of literature in a limited way. The other extreme is to introduce literature in various forms from the very beginning, whether it be in Mother Goose tales, poems put to song, or children's plays. Povey (1967) pointed out on this last point, that in the U.S. from the beginning, grammar has been separated from literature in our curricula (as is attested to by separate reading and literature classes from elementary through to highschool) which shows that we put a heavy emphasis on literature from the beginning. He continues by saying, "that if educators here, have seen advantages in allowing language and literature to interact, why not do so in ESL, almost from beginning levels."

Now this whole question of vocabulary difficulty has a certain amount of relativity. Do native speakers ever have difficulty with vocabulary, literary or otherwise? Obviously they do, and in varying degrees. Is anything wrong with having a little bit of difficulty in understanding words, or in understanding what certain novel arrangements of known words mean? Again I think not. It is simply a question of degree. In real language use situations, will a learner -- or native speaker for that matter -- ever know and understand all the words that are spoken? Perhaps. But don't we usually listen and read for meaning with some disregard for the total quantity of individual linguistic units? My contention is that we do. So if in real life we will have to scramble for meaning among the floods of sounds and gestures that come pouring forth, shouldn't we practice this type of activity in class? Again I would answer, yes. Finally then couldn't it be seen as valuable to a communicative language approach, to learn how to read for understanding by looking in a dictionary when absolutely necessary, but otherwise guessing and making inferences from context and situation? I would say yes, because the skills of understanding meaning from context are very similar in both listening and speaking.

Povey (1967) states that "the existence of a 'recognition' vocabulary is well known. There can be a general comprehension even when there has

not been a precise understanding of a certain syntactic structure." Marckwardt, in dealing with the large quantity of English vocabulary (500,000+), feels that of the three ways of learning vocabulary, memorizing lists, using dictionaries, and learning from context, the last is the most effective for the English language. He points out that since native language speakers learn more by context, especially in the realm of receptive vocabulary, why shouldn't ESL students do so also. Regarding fullness of meaning he says that "most of us acquire new words in our receptive vocabularies by adducing their meanings from the contexts in which they occur. Repeated exposure, where the word appears in slightly different contexts, fleshes out its meaning and also prepares one to use it if the opportunity presents itself."

Although the question of vocabulary is an important one in choosing which works to read, there is no need to be overly concerned about finding absolutely easy-reading books. My German teacher always told me to read for meaning, and not to look up every word. At first it was very hard, so much so that I occasionally was tempted to throw fits. But after the initial weeks of learning this new skill (Isn't all learning painful at times?) I was able to read with reasonable speed and understand the majority of what I was reading. And it still works, even when I'm reading difficult passages in English, my native language. Contextual vocabulary learning is a skill all L2 learners need, and will be useful in improving both reading and listening skills, in the second and native languages.

SPEAKING, READING, LISTENING, AND WRITING

Regarding the four skills of reading, writing, speaking, and listening, let me briefly comment on how they can be benefited and improved through the use of literature in the classroom. As for speaking skills, I must say that in all of my conversation classes, one of the main difficulties has always been to strike up a conversation. What is so easy beside a fireplace, or over a cup of coffee in a cafe, is often the bane of the language teacher. "I can't even get these kids to talk, so how do I know if they have learned anything!?" The personal involvement and emotional appeal which literature can provide is a ready made source of material for conversations about any number of subjects. It serves then as a springboard into a wide range of discussions involving vocabulary, structures, situations, historical periods, cultural traits or societal norms that the teacher may desire to deal with.

Dramatic readings of poems, scenes from plays, or famous narrative passages, are another excellent means for practicing pronunciation, intonation, rhythm, pitch, and fluency in general. The possibility of students giving oral reports, is yet another means of providing a meaningful situation for communicating. The student can become a master of a certain subject, and then figure out how he is going to express himself, fluently and meaningfully, before the whole class. There are just three ways in which I have seen the use of literature to be helpful in improving my overall feel for a language and fluency in oral expression and communication.

Aural comprehension can be aided by listening to taped versions of poems and songs, or by listening to dictations taken from the text and read aloud by the teacher (story time), or in more advanced classes by watching video productions of famous literary works such as Oliver Twist or My Fair Lady. The emotional involvement of literature is still there, and the aural comprehension will be greatly increased by seeing the speakers' faces, and by first having read the whole, or parts of the whole work beforehand. In Pamplona, Spain, where I taught ESL for two years, there is

an English Academy called the "English Film Center". They teach English by varying methods, but the central link in the academy is that once a week full length movies are shown on video in English for all students to see. Students can see the videos more than once if they have time to come back for a second session. Before the films are shown each student reads a copy of the play or novel which the movie is based upon. (These are sometimes simplified or abridged versions.) This then provides them with multiple exposure to excellent language and culture models, anything from "The Great Gatsby" to "Ghandi" or "Casablanca", after which they feel they have really experienced and understood ideas, emotions, and life in the target language culture, in this case English. This is one of the more innovative approaches I have come upon for using literature in a pleasing and viable way on a large scale.

As far as reading is concerned I have already dealt with its use in vocabulary building. The very exposure to language of different levels, registers, and social milieu, that reading nessecitates, is a benefit in itself. If one really considers it to be a form of interaction between the author and the reader, then this is real communication, and it is very little different from sitting down for a cup of coffee with a friend and talking to him in the target language for hours. The advantage is that you can stop and have your friend repeat exactly what he said (reread the passage) and have him talk as slowly as you need him to at first. Books are more portable than friends, and why not let an author who has lived, studied, worked, and thought about his country just sit down and tell you about it? Perhaps not all literature can be treated as personally as this, yet with practice one learns to pick out the works that discuss one's personal interests.

Writing also improves, through the practice gained in creating themes for the teacher, or through writing reactions to literature in personal journals as Spack (1985) requires students to do in her classes. This last activity, for Spack, is a low anxiety way of inducing students to write creatively for expression of ideas without having to pay so much attention at first to form. She has had very good results with journal writing, because of the high amount of affective motivation available when a literary work has been effectively dealt with in class. To feel the flame of inspiration burning deep down inside, even and especially in the target language, is a wonderful reassurance that actual understanding and communication has taken place. A student's self-confidence and self-esteem are greatly boosted when he finds that he can read, think, analyze, and then write effectively in the target language. McConochie (1985), Spack (1985), and Zamel (1982) all deal with methods of structuring and organizing the writing process in ESL teaching in such a way as to interest students, and help them learn about the writing process as a way of discovering meaning. This has obvious value in the language learning process, from thinking in the TL to expressing thoughts concisely and logically on paper. The crossover into speaking skills is very rewarding for those who practice.

SOCIOCULTURAL COMPETENCE

Now regarding sociocultural competence, I think that many of the new methods are stressing that for true communicative competence, a student must obtain a high level of sociocultural competence as well. If either the linguistic or cultural components of a language break down, there is a break in communication, as can be attested to by the American business man who has never been hugged or kissed in his life, by a man, and is immediately set on edge by this very occurrence at his first informal meeting in the home of some company friends. Perhaps this same businessman is chased all around the office by a Latin colleague who insists on getting

right in his face (garlic breath and all) to talk about the all important multimillion dollar deal. The more the American backs up, the more the Latino advances, both feeling the other is just a little rude without perhaps knowing exactly what motivates the other's strange and abnormal behavior. In both cases very little will be communicated, either by the American or to the American (although a very active inner dialog will be taking place in both men's heads), due to the cultural noise created by their lack of sociocultural awareness. In fact it could, in extreme cases, lead to very hard feelings, erroneous accusations, and the loss of a business deal. In less extreme cases it often leads at least to a constant semiconscious level of anxiety and discomfort, which in no wise aids the comprehension of the linguistic side of the communication.

Brown (1987) in his book on language learning and teaching, mentions four factors as the major sociocultural hurdles in language learning. They are stereotypes, attitudes, acculturation, and social distance. In another paper I have dealt with them, but would like to reiterate here that I think literature, if used appropriately, can be a prime tool for overcoming these hurdles. Brown said that in many cases the first two, sterotypes and acculturation, can be overcome simply by providing exposure to the reality of the target language, its speakers and culture. This serves to dispel the false ideas and impressions gained during childhood, from parents, peers, and bad experiences. For acculturation and social distance problems, exposure again, empathy building experiences, and extended contact with the language are what is needed.

Of course there are many ways of providing this much needed exposure. But I would begin by arguing the point of expense in travelling to the TL country or communities in the U.S., that many language learners are faced with. Literature is a relatively inexpensive way of coming into fairly full emotional and cerebral contact with a foreign culture. Before you are chased around the room by a friendly Latino businessman, why don't you read about it and confront the cultural conflict head on in a class discussion on a novel, play, or video production. Literature should not be read only for culture learning, for this would cheapen the literary experience, as Spack points out, but why not show the real people using the language in uncontrived, natural dialogs as literary authors can do best? (It's better yet if you can see it take place on the screen.) My point is that a certain amount of crosscultural frustration can be avoided by explicit teaching of cultural differences in language classes before actual use of the language is needed (even during use in the case of in-country language learning). Literature is a prime tool for this purpose.

McCleod (page 539 in Croft 1980) says that culture should be taught explicitly in the language classroom because students do not have enough time to learn it implicitly as native speakers do. The point is well taken, for one cannot understand foreign cultures until one has analyzed one's own culture first, which few native speakers can do without help from the anthropologists, because they have learned their own culture implicitly. If you can't deal explicitly with your own culture, you won't be able to deal explicitly or implicitly with another culture in any short period of time. McCleod goes on to say that each language learner should be taught to be an anthropologist, for only a mind trained in spotting and sensing cultural differences, can learn them and comprehend their significance for communicating effectively.

Marckwardt (1978, p. 17) agrees with this need for explicit culture teaching and points out that, "Literature can be made to fulfill a cultural objective if careful plans are made toward this end, involving the choice

of literary selections and manner of presentation. Needless to say, a set of academic lectures on the topic will fail of their purpose; here we are in the realm of the affective and not the cognitive." This last statement calls to mind what was mentioned earlier about how literature can strike a chord in a person, producing a positive emotional response, that can be beneficial in bridging the empathy gap which so often exists between cultures.

Marckwardt (p. 49) goes on to say that "cultural awareness, sensitivity, and rapprochement are not automatic by-products of foreign language instruction or of exposure to a foreign literature. If they are to be achieved at all, they must be planned for and built into the course of study." Of course many strategies for teaching culture explicitly have been devised and put into practice. But more than just anthropological readings and lectures on culture, Marckwardt feels, and I agree with him, that "the most effective presentation will be through observing people in contexts, in situations where they act and react to each other in terms of their culture." Again as in the structure and vocabulary discussions, we have contextual learning, this time in a culture specific situation. Culture is taught by observing it in action.

Some might argue for using simple dialogs and other contrived teacher texts to teach cultural points since literature can be so difficult. But Spack (1985) points out that it makes no sense to debunk literature as irrelevant, and then create one's own cheap dialogs. For text dialog, she prefers using literary excerpts for their "context meaningfulness and excellence of language", as well as uncontrivedness.

Marckwardt (p. 50) continues his discussion on the matter of culture by saying that one interesting and unique factor in using literature is that "Characters in literature, [reveal] not merely their actions, but their thoughts and emotions as well, in a form which permits repeated impact and careful analysis." In everyday life many native speakers do not open up to strangers, and it is not impossible to find people who have never had the opportunity to sit down for a heart to heart talk with a native speaker, and hence have never really gotten inside the other's head to figure out why he acts, thinks, or reacts as he does. Similarly, Collie (1987, p. 3) states that in literature "a reader can discover their (the literary characters') thoughts, feelings, customs, possessions; what they buy, believe in, fear, enjoy; how they speak and behave behind closed doors." Is it not a coveted experience for all language learners to find that perfect host family, who will adopt, teach, feed, explain, and ease the learner through culture shock? Here, then in literature, a beginning can be made by entering into a foreign family's home, mind, and heart, and observing and hearing about all the intimate details of family life, which can provide so many clues to the cultural riddles set before the language learner. As Collie says, (p.6) "Fiction: summons the whole person into its own world." If this is true, and I have found it so for well written works, then foreign language friends can be made without ever leaving the armchair.

DRAMA FOR LANGUAGE AND CULTURE LEARNING

One other very unique use to which literature is being put, is that of drama. Oller (1983) in his book, includes some good articles on the use of drama in language teaching, and why it is successful, so I won't go into all that. Role plays, skits, and scenarios are very helpful in achieving certain objectives. But producing a full-scale drama (perhaps only one-act plays) can be an extremely rewarding experience, in culture learning, oral expression, vocabulary building and much more. Susan Stern in Oller (1983) talks about using drama, but one of the most unique experi-

ences I have heard about, was one that is used by Central College of Iowa's Vienna study-abroad program.

They happen to have a very gifted man working with them who is skilled in drama, loves literature, music, and art, and also happens to know half of the important people in the city of Vienna. Well, usually every Spring, he picks a short play, usually a Viennese period piece, with authentic language exchanges, cultural peculiarities (isn't that what comedies often portray?) and a little bit of love and excitment. All students are eligible for tryouts, but those who get the lead parts know that they have their work especially cut out for them. For the linguistic aspect of the production, tapes are made by native speakers of each part, including the dramatic nuances needed to create the effects desired by the director. Scripts are memorized, and accent, pitch, tone, speed and any other necessary items are dealt with through the cassette lab work, and through intensive sessions with the director.

On stage, blocking is done, and Austrian patterns of cultural interaction are discussed, demonstrated, practiced, and then incorporated in the students life. Costumes, foods, furniture, ideas, and pastimes are studied, discussed and then seen in action and on stage. The director happens to have an "in" with the National Theatre, so authentic and often very expensive costumes are used. Of course the involvement has to be total, and even those simply helping with props, make-up, or costumes, have heard the lines over and over and have been so closely involved with the production that they feel as if they are part of Vienna. The high degree of integrative motivation which this produces does wonders for the whole production. The plays are usually performed at a local high-school, and all the friends and acquaintances of the students are invited. The comments I heard from the director, other teachers, and program advisors, indicated that an incredible degree of accuracy, fluency, purity of accent, and acting skill would inevitably result from the group almost every year.

One teacher said that the sensitivity with which the actors and actresses portrayed the roles was comparable with that of some professional performances she had seen. The director himself said that more often than not, through the intensive linguistic work, practice, and dedication, the Austrian audience could not detect if certain roles were played by Americans or nationals. Of course there would always be some who for all the work could not attain this high level of accent purity, but even they were thrilled about the increased ability they had to feel at home in the language. Culturally, the drama experience helped them bridge the gap that existed between Vienna and the good old U.S.A. The overall impression that I received was that the students so thoroughly did their homework, and practiced their parts, that they really felt, looked, spoke, and acted more like Austrians, both at home in the dorms, and at dinner engagements; in class and in the park; in cafes and in concerts. In short wherever they went in the city they felt like one of the Viennese. Of course the Viennese could still spot them in a crowd, but part of the need in language teaching and learning is self-esteem and self-confidence, which these drama productions provided.

CONCLUSION

My main point here, has been to show that literature has many unique applications which are just beginning to be touched on. Gone are the days when word for word translation of old stodgy texts sufficed for what is called foreign language learning. So let's clean out the negative attitudes that are held toward using literature, and come up with any number of new and unique ways that it can be used in teaching. Of course it is not a cure all, and I would not intend to suggest that, but it is

safe to say that literature provides us with a ready supply of quality language models, on every imaginable subject area, era, situation, conflict, and social class.

Depending on what literature is chosen for which group, the difficulty factor can be kept within logical bounds, and interest can be sparked in the readers. Commentaries on famous works abound, and books and articles on how to teach literature effectively can be had by teachers all over the world to improve their teaching skills. Approaching the work based upon aesthetic appeal, and meaning and content will serve to interest students, and stir up their ideas, while inducing very full contact with the work, the language and the culture of the target language. In short the experience with literature can and should be enjoyable and productive in cerebral, emotional, as well as linguistic activity. If all of this happens, I am sure that literature will take its rightful place as an old tool used effectively in the new communicative methods.

So in learning culture and language, why not let the authors who have pondered so deeply on their country speak for themselves? They will tell the language learner all about their culture, families, values, habits, songs, foods, clothing, morals, and moods. They can be argued with in class discussions or in writing, and can be questioned as to their personal conclusions on certain values or vices. As for variety, you can read of nobility, you can read of urchins; you can read of coal miners, you can read of criminals; you can read of soldiers, you can read of lovers; you can read of mothers, you can read of children. Literature, in one way, is the sum of a country's language and culture, ordered, pondered, and expounded in small volumes for anyone to learn at their leisure, slowly but surely. From beginning classes through advanced, this is valuable, and can help make foreign language learners functional if they practice what they know and learn.

REFERENCES

Brown, H. Douglas Principles of Language Learning and Teaching, 2nd Ed. Englewood Cliffs: Prentice Hall, 1987.

Collie, Joanne and Slater, Stephen. Literature in the Language Classroom. Avon: Cambridge University Press, 1987.

Croft, Kenneth. Readings on English as a Second Language. Boston: Little, Brown and Company, 1980

Marckwardt, Albert H. The Place of Literature in the Teaching of English as a Second or Foreign Language. University of Hawaii Press, 1978.

Marshall, Madeline. "Love and Death in Eden: Teaching English Literature to ESL Students." TESOL Quarterly 13 (1979) Page 331.

McConochie, Jean. "'Musing on the Lamp Flame': Teaching a Narrative Poem in a College-Level ESOL Class." TESOL Quarterly, Vol. 19, No. 1, March 1985.

McKay, Sandra. "Literature in the ESL Classroom." TESOL Quarterly, Vol. 16, No. 4, Dec. 1982, (Page 529)

Oller, John W. Jr. Methods That Work. Cambridge, MA: Newbury House Publishers, 1983.

Povery, John F. "Literature in TESL Programs: The Language and the Culture." TESOL Quarterly, Vol. 1, No. 2, 1967. (Page 40.)

Space, Ruth. "Literature, Reading, Writing and ESL: Bridging the Gaps." TESOL Quarterly, Vol. 19, No.4, Dec. 1985.

Tucker, Richard and Gatbonton, Elizabeth C. "Culture Orientation and the Study of Foreign Literature." TESOL Quarterly, Vol 5, No. 2, 1971, (Page 137)
Zamel, Vivian. "Writing: The process of Discovering Meaning." TESOL Quarterly, Vol. 16, No. 2, June 1982, (Page 195)

TESTING FOR LANGUAGE PROFICIENCY

COMMUNICATIVE AND COGNITIVE LANGUAGE PROFICIENCY:

IMPLICATIONS FOR TESTING IN THE L2 CLASSROOM

Marguerite G. MacDonald

Wright State University
Dayton, OH

Two essentially different linguistic skills are involved in the mastery of a second foreign language. Stephen Krashen differentiates these two abilities through the labels acquisition and learning. Krashen (1982:10) maintains that "adults have two distinct and independent ways of developing competence in a second language." The first way is through acquisition, which parallels first language acquisition in the child. Acquisition is a subconscious process, using language for communication. The second way is through learning. This is the conscious knowledge of language rules. In this paper I will explore the implications of this dichotomy, that is, the difference between communicative and cognitive language proficiency, for testing in the second language classroom.

Krashen's (1982) claims for acquisition and learning are closely related to his Monitor Hypothesis. The Monitor Hypothesis proposes that acquisition and learning are used in different ways. Acquisition initiates utterances in a second language and is responsible for fluency. In contrast, learning is used as a monitor or editor.

Others have likewise distinguished two separate areas of language proficiency. Herandez-Chavez et al. (1978) point out the difference between natural communication and linguistic manipulation. Similarly, Brown (1980) refers to two types of learning, correlated with field dependence and field independence. He notes that

> one kind of learning implies natural, face-to-face communication, the kind of communication that occurs too rarely in the average language classroom. The second kind of learning involves the familiar classroom activities: drills, exercises, tests, and so forth. (Brown 1980:92)

A similar dichotomy of language proficiency has been proposed by Jim Cummins (1980:175), who claims that

> a dimension of cognitive/academic language proficiency can be empirically distinguished from interpersonal communicative skills such as accent, oral fluency and sociolinguistic competence in both first and second languages (L1 and L2), and that cognitive/academic proficiencies in both L1 and L2 are manifestations of the same underlying dimension.

Cummins (1980) maintains that cognitive language ability is related to both general cognitive skills and academic achievement and that it can be assessed by a variety of reading, writing, listening, and speaking tests. However, while cognitive language proficiency is closely related to IQ and other aspects of academic achievement, Cummins (1980) claims that communicative skills are not. He points out that "with the exception of severely retarded and autistic children, everybody acquires basic interpersonal communicative skills in a first language regardless of IQ or academic attitude" (Cummins 1980:176).

This claim does not, however, hold true for second or foreign language learning. Unlike cognitive language skills, which Cummins (1980) maintains are easily transferrable from the native to the target language, communicative skills in the native language are no guarantee of communicative ability in the target language. Proficiency in basis communicative skills may be outside the reach of the language student, particularly in the foreign language classroom.

Traditionally, the classroom has been the accepted domain for the development of cognitive processes. Learning is associated with the formal environment, with those abilities related to IQ and academic attitude. In contrast, communicative skills, both in the first and target language, have for the most part been acquired in an informal, natural setting. Thus in the acquisition of a foreign language, there is the danger that communicative skills will be neglected almost entirely, since the informal setting does not provide the necessary exposure to the target language.

Unless there is adequate exposure to the target language, communicative skills cannot be acquired. Krashen (1982) further qualifies exposure, claiming that the input must be comprehensible. Recent methodology in second and foreign language teaching has recognized the need to provide comprehensible input in order to develop communicative skills, and few language teachers would deny the importance of these skills. Nevertheless, current work in Proficiency-based language teaching has reemphasized the role of cognitive grammatical skills, to some extent at the expense of communication.[1] Omaggio (1986:49) justifies this position by claiming that "native speakers' reactions to learners' errors revealed that lexical and grammatical errors are the most obstructive to communication." She also cites Ensz' (1982) study, which found communication without grammatical accuracy to be less acceptable than less communicative, grammatically accurate production.

Higgs and Clifford (1982) likewise support an emphasis on cognitive grammatical skills over communicative skills, claiming that emphasizing communicative activities too early will cause fossilization in the speaker at lower levels of communication (2 to 2+ on the Foreign Service Institute interview rating scale).[2] They warn that

> there appears to be a real danger of leading the students too rapidly into the "creative aspects of language use," in that if successful communication is encouraged and rewarded for its own sake, the effect seems to be one of rewarding at the same time the incorrect communication strategies seized upon in attempting to deal with the communication situations presented. (Higgs and Clifford 1982: 74)

Higgs and Clifford (1982) maintain that if students are to reach level 3 on the Foreign Service Institute (FSI) exam, then cognitive grammatical skills must be an important part of the curriculum, with pronunciation

---

[1]See notes on page 151.

and fluency receiving proportionately less attention.[3] Although Higgs and Clifford (1982) clearly recognize that communicative ability is necessary, their emphasis on grammatical accuracy reduces the centrality of pronunciation and fluency. This position has been interpreted as a justification for avoiding communicative activities and returning to discrete point grammatical exercises, especially in the early stages of language learning.

However, the evidence supporting the position held by Omaggio (1986) and Higgs and Clifford (1982) is questionable. In a recent study by MacDonald et al. (1986) grammar proved to be the least significant of the four areas measured in relation to communicative ability. Using the SPEAK test, the institutional version of the Test of Spoken English (TSE), the investigators established the correlation of pronunciation, fluency, grammar, and comprehensibility with the ability of the subject to be understood by native speakers of English. The comprehensibility scores of the subjects ranged from 170 to 300 on the SPEAK test, equivalent to 2 through 3+ and above on the FSI scale.[4] As Table I shows, the comprehensibility score was the most highly correlated with the subject's ability to be understood. However, the pronunciation score had a comparable correlation. Fluency was less highly correlated, while grammar showed a significantly lower correlation.

Further, while the intercorrelations for the four scores were high in all instances, as Table II indicates, the correlation of grammar with the other three factors was consistently lower than the intercorrelations of the remaining three scores.

The results obtained by MacDonald et al. (1986) parallel those found by Clark and Swinton (1980) for their study on the TSE. Clark and Swinton (1980) also found that the grammar score of the TSE was more highly corre-

Table I
Correlation Between SPEAK Scores
and Ability to be Understood
(Pearson's Correlation Coefficients)

SPEAK Score	Ability to Be Understood
Comprehensibility	.48
Pronounciation	.47
Fluency	.42
Grammar	.35

Table II
Intercorrelation Between SPEAK Scores
(Pearson Correlation Coefficients)

	Pron.	Fluency	Gram.
Comprehensibility	.90	.92	.84
Pronunciation		.86	.80
Fluency			.84

lated with the subjects' TOEFL scores than were the remaining three TSE scores.[5] This demonstrates that the grammar score more closely resembles the cognitive abilities measured by the TOEFL exam.

Therefore, grammar is not the central factor for proficiency in communicative skills, even when speakers have achieved levels of 3 or 3+. Further, as Higgs and Clifford (1982) admit, a heavy emphasis on grammatical accuracy produces students less able to communicate in the language after four semesters, the maximum period of time normally devoted by college students to the learning of a foreign language.

Another danger in the position held by Higgs and Clifford (1982) and Omaggio (1986) is that it implies cognitive grammatical ability will lead to superior communicative skills. Higgs and Clifford (1982:78) claim that

> if students develop both receptive and productive competence at a pace that allows for reasonable internalization--albeit under heavy monitoring at first--then utimately it will be easier for them to activate skills that are measurably present, though passively, than it will be for them to dismantle reinforced or fossilized skills and replace them with different ones.

Yet cognitive ability does not necessarily convert to communicative skill, supporting to some extent Krashen's (1982) claim that learning does not become acquisition. Studies on age and language acquisition bear out this assertion. These studies have produced seemingly contradictory results. In general, short-term language exposure studies have favored older children and adults over younger children in their target language ability, whereas long-term language exposure studies have favored those who began acquisition at a younger age.

This apparent contradiction was explained by Cummins (1980), who noted that the conflicting results could be attributed to the different types of measurement employed in the studies. He claimed that adults did better on tests measuring cognitive proficiency but that this was not the case for fluency and pronunciation. He concluded that

> a cautious generalization from these findings is that oral fluency and accent are the areas where older learners most often do not show an advantage over younger learners. . . . given sufficient exposure to the L2 and motivation to learn L2, older learners will perform better than younger learners on any measure that loads on a CALP [cognitive/academic language proficiency] factor. (Cummins 1980: 180)

However, this cognitive factor does not convert into communicative skills. Were this the case, adults would, in the long run, remain ahead of their younger counterparts. In contrast, long-term studies show that children surpass adults in communicative skills. It appears then that, while adults and older children have superior cognitive abilities, younger children are better able to acquire communicative skills.

In the studies that measure long-term exposure to a second language, the methodology has addressed a speech that is more natural and less monitored, that is, a speech that emphasizes communicative skills. Asher and Garcia (1969), in their long-term study on pronunciation, required their subjects to practice sentences and say them after they were comfortable with them. When the sentences were actually produced by the subjects, there was no model immediately preceding the utterance. While not communicative, the speech was more natural than that produced by immediate repetition.

Seliger, Krashen, and Ladefoged (1975) used self-evaluations for their study on long-term acquisition. These judgments were concerned with how native the speakers considered their pronunciation to be. Thus communicative skills were evaluated.

In Oyama's (1976) study the methodology relied on a holistic evaluation of pronunciation evaluated by graduate students in linguistics. The speakers both read and recalled material. Once more, particularly for the recall exercise, the speech production was communicative in nature. A second study by Oyama (1978) measured the subjects' ability to comprehend natural language when there was interference. Again, communicative language was employed, in this case in a receptive capacity.

Patkowski (1980) examined the syntax of oral interviews patterned after the FSI interview. Unlike most measurements of syntax, the task was not focused and did not call attention to discrete points.

MacDonald (1986) used an informal interview procedure to determine phonological and nonphonological variation in the second language of high school seniors who had acquired the target language as preadolescents. She found that those who had begun acquisition at an earlier age showed significantly less variation in both pronunciation and grammar.

In contrast, the short-term studies primarily involved focused, monitored abilities that, as Cummins (1980) noted, tap cognitive skills. The short-term studies comparing children and adult language ability generally preferred repetition for pronunciation measures and translation or controlled examinations, such as a Berko (1958) completion test, for measures of morphology and syntax.

In a study by Asher and Price (1967) short commands were followed. Although this was a more natural use of language, the abbreviated nature of the message forced a focus on monitoring rather than continuous speech.

Olson and Samuels (1973) and Snow and Hoefnagel-Höhle (1977) relied on repetition of words and short utterances in their studies on pronunciation and age. Olson and Samuels (1973) trained their subjects by using drills over a period of time, while Snow and Hoefnagel-Höhle (1977) employed a drill exercise immediately preceding the testing. Both approaches focused on the task rather than natural language production.

In a second study, Snow and Hoefnagel-Höhle (1978) used several measurements, including repetition, minimal pair discrimination, completion, translation, grammatical judgment, vocabulary testing, story comprehension, and story retelling, in order to evaluate various aspects of language comprehension and production. They tested their subjects on three occasions at intervals of several months. Initially, the older subjects were more successful at all tasks. This is not surprising since early second or foreign language production usually relies heavily on cognitive ability through monitoring. However, by the third testing the younger subjects began to take the lead in communicative skills. While the older subjects continued to excel at the structured tasks, for the two exercises that most closely approached a natural speech context, story comprehension and story retelling, the 6- to 7-year-old group surpassed the other groups. Although the 3- to 5-year-old group did not show comparable gains, it can be argued that this group was still in the process of first language acquisition.

For short-term studies comparing younger and older children, the same types of measurement have been used as for those studies of adults and children. To test morphology and syntax Fathman (1975) relied on the completion method, again inspired by Berko's (1958) research. Not sur-

prisingly, the older children did better on the controlled test. However, the younger children were judged superior on pronunciation. Unlike the controlled syntax and morphology tasks, the pronunciation measurement was contained in the natural context of answering a question.

Ekstrand (1976) used repetition for his pronunciation measurement. To test other aspects of language proficiency he employed tests for listening comprehension, reading comprehension, free oral production, and free written production. While the older children excelled in the more controlled tasks, free oral production proved not to be related to age. Listening comprehension and written production were dependent on length of residence rather than age. Another study by Ekstrand (1978) used immediate repetition and translation. Again, on these controlled tasks the older children excelled.

In general then, when the task was focused, allowing monitoring and relying on cognitive language skills, the adults and older children were superior. However, when enough time was allowed for communicative skills to develop and begin to supplant reliance on cognitive monitoring, the younger children proved more proficient at natural language measurements. Thus Higgs and Clifford's (1982) suggestion that cognitive grammatical ability will lead to communicative ability is not supported by the evidence. If cognitive ability could be converted into basic communicative skills, then adults and older children would clearly surpass younger children in the long-term studies as well.

Since communicative skills are qualitatively different from cognitive language proficiency, it is necessary to test both types of language ability in order to obtain an understanding of overall proficiency. However, teachers rarely measure the communicative aspect of language. Even if oral language is involved, the task is usually focused, allowing time for monitoring so that cognitive language ability, not communicative skills, is evaluated.

There are several reasons for this approach to language testing. One reason is that, as mentioned previously, the domain of the classroom has traditionally been cognitive learning. It is the area correlated to IQ and academic achievement. Further, it is the area most easily taught, over which the instructor feels the greatest control. Cognitive language proficiency is also the most easily measured, since discrete point tests can be used. Because written skills parallel oral production for cognitive proficiency, teachers can rely on the more easily graded written medium. Finally, if the instructor believes that this cognitive ability can be converted into communicative skills, then there appears no reason to assess these elusive communicative skills separately.

Measuring communicative skills has been viewed as a time-consuming process. The teacher must interview each student separately, which requires several days of classroom time. To evaluate this production also presents a difficulty in control because of the less structured, more individual nature of communicative language. Thus a conflict arises. In order to test efficiently and accurately, a controlled measurement is preferred. Yet the nature of the control may eliminate any possibility of measuring communicative skills.

Once, however, teachers recognize the need to separately assess communicative skills, then they must address the manner of doing so efficiently. One solution is to evaluate one or two students during each class session. So as not to focus on the measurement, the teacher can do this testing as part of the normal interaction in the classroom. If a communicative session occurs two or more times during each class period, the instructor can holistically assess the last one or two students in

each session, determining their level of pronunciation, fluency, and general communicative ability within the sociolinguistic context. The teacher can then record these results while gathering materials for more structured activities or after dismissing the class at the end of the period. This approach reduces the possibility of the students becoming intimidated by the testing procedure and attempting to heavily monitor their output.

Because cognitive ability and communicative skills are two separate aspects of language proficiency, each must be considered separately in any evaluation process of language proficiency. By comparing the student's evaluations for communicative skills throughout the semester or year, the instructor can establish the student's progress in achieving communicative proficiency. When joined with the cognitive language measures, these scores will more accurately reflect the overall language proficiency of the student.

NOTES

[1]The division between communicative and cognitive grammatical skills is complex. Through monitoring, both skills may operate simultaneously. Particularly in early language acquisition, speakers may rely on heavy cognitive monitoring to compensate for their lack of communicative skills.

[2]The 2 to 2+ speaker is "able to fully participate in casual conversations, can express facts, give instructions, describe, report on, and provide narration about current, past, and future activities. [The speaker can] discuss concrete topics such as own background, family, and interests, work, travel, and current events. [The speaker is] understandable to an NS [native speaker] not used to dealing with foreigners [but] sometimes miscommunicates" (Higgs and Clifford 1982:63).

[3]The 3 to 3+ speaker "can converse in formal and informal situations, resolve problem situations, deal with unfamiliar topics, provide explanations, describe in opinions, and hypothesize. [The speaker can discuss] practical, social, professional, and abstract topics, particular interests, and special fields of competence. Errors never interfere with understanding and rarely disturb the native speaker (NS). Only sporadic errors in basic structures" are produced (Higgs and Clifford 1982:63).

[4]The Educational Testing Service (1982:19) explains that when correlating the TSE scores with those of the FSI interview "above level 3+, a ceiling effect begins to appear." Therefore, although some of those with scores above 260 may be in the 4 range, the TSE cannot discriminate between 3+ and higher levels.

[5]Clark and Swinton (1980:19) show the following correlations between the components of the TSE and the TOEFL exam: grammar .70, Fluency .60, comprehension .57, and pronunciation .56.

REFERENCES

Asher, James, and Ramiro Garcia. 1969. The optimal age to learn a foreign language. Modern Language Journal 53.334-341.

Asher, James, and Ben Price. 1967. The learning strategy of Total Physical Response: some age differences. Child Development 38.1219-1227.

Berko, Jean. 1958. The child's learning of English morphology. Word 14.150-177.

Brown, H. Douglas. 1980. Principles of language learning and teaching. Englewood Cliffs, N.J.: Prentice-Hall.

Clark, John, and Spencer Swinton. 1980. The Test of Spoken English as a measure of communicative ability in English-medium instructional settings. Princeton, N.J.: Educational Testing Service.
Cummins, Jim. 1980. The cross-lingual dimensions of language proficiency. TESOL 14.175-187.
Educational Testing Service. 1982. Test of Spoken English manual for score users. Princeton, N.J.: Educational Testing Service.
Ekstrand, Lars. 1976. Age and length of residence as variables related to the adjustment of migrant children with special reference to second language learning. Proceedings of the Fourth International Congress of Applied Linguistics, ed. by G. Nickel, 179-197. Stuttgart: Hochschulverlag.
_____. 1978. English without a book revisited: The effect of age on second language acquisition in a formal setting. Didakometry 60. Malmo: Department of Educational and Psychological Research, School of Education.
Ensz, Kathleen. 1982. French attitudes toward speech errors. Modern Language Journal 66.133-139.
Fathman, Ann. 1975. The relationship between age and second language productive ability. Language Learning 25.245-253.
Hernandez-Chavez, Eduardo, Marina Burt, and Heidi Dulay. 1978. Language dominance and proficiency testing: some general considerations. NABE Journal 3,41-54.
Higgs, Theodore, and Ray Clifford. 1982. The push toward communication. In Curriculum, Competence, and the Foreign Language Teacher, Ed. by Theodore Higgs, 57-79. Lincolnwood, Il.: National Textbook Company.
Krashen, Stephen. 1982. Principles and practice in second language acquisition. Oxford: Pergamon.
MacDonald, Marguerite. 1986. The relationship between sociodemographic factors and linguistic variation in the Cuban Community. Paper presented at the Fifteenth Annual International Bilingual/Bicultural Education Conference. Chicago, Illinois.
_____, Henry Limouze, and Ann Keucher. 1986. The relationship between SPEAK scores and student evaluations. Paper presented at the Third Annual LAU Center Conference and Ohio TESOL Annual Fall Conference. Columbus, Ohio.
Olson, Linda, and S. Jay Samuels. 1973. The relationship between age and accuracy of foreign language pronounciation. Journal of Educational Research 66.263-267.
Omaggio, Alice C. 1986. Teaching language in context. Boston: Heinle and Heinle.
Oyama, Susan. 1976. A sensitive period for the acquisition of a non-native phonological system. Journal of Psycholinguistic Research 5.261-285.
_____. 1978. The sensitive period and comprehension of speech. Working Papers on Bilingualism 16.1-17.
Patkowski, Mark. 1980. The sensitive period for the acquisition of syntax in a second language. Language Learning 30.449-472.
Seliger, Herbert, Stephen Krashen, and Peter Ladefoged. 1975. Maturational constraints in the acquisition of second languages. Language Sciences 38.20-22.
Snow, Catherine, and Marian Hoefnagel-Höhle. 1977. Age differences in the pronunciation of foreign sounds. Language and Speech 20.357-365.
_____. 1978. The critical period for language acquisition: evidence from second language learning. Child Development 49.1114-1128.

THE PRAGMATIC DEMANDS OF PLACEMENT TESTING

Francis J. Sullivan, Jr.

Temple University
Philadelphia, PA

The study I am reporting on today examines how the overall ranking of students' placement-test essays was influenced by the pragmatic form of the texts the students had produced. Specifically, it documents how patterns in students' use of information assumed to be either old, inferrable, or new for readers affected the evaluation of the texts and suggests that these effects are, at least partly, a function of the fact that this was a test--a demand that writers not only communicate but also display their ability to do so.

The study used essays written by 99 randomly selected lower-division students entering Temple University during the summer of 1982. Non-native students were deliberately excluded from the sample. All students had written essays that responded to the same rhetorical demands--a trans-actional piece of prose explaining the writer's position on a current public issue. The essays had been ranked using a modified holistic scoring system based on that suggested by Diederich (1974) and Cooper (1977).

The analysis of the texts was based on a taxonomy of given/new information developed by Prince (1981), which classifies information represented by noun phrases according to how familiar readers are assumed to be with it. Prince offers three categories of what she terms "Assumed Familiarity." Overall, information is assumed to be either New, Inferrable, or Old (which she terms "Evoked") for readers. Definitions for all the terms are given in the appendix.

Anytime I introduce information into a discourse, written or spoken, it is assumed to be New. But, the distinction Prince makes between Brand-new and Unused information is significant. The former type represents things I assume are new in both the text and in your head. I haven't mentioned the entity before, and I assume you haven't heard of it before either. This kind of information is cast in indefinite noun phrases. The latter, however, represents things that may be new to the text but I assume are old in your head. You may not be thinking about that entity right at the moment I introduce it, but I have reason to believe you have the information available to you in memory. Proper names make up the bulk of such entities--but not all. For instance, if I had begun this paper with the example given in the appendix under the category Unused, I am confident that all of you would have known just what world I was referring to, despite the fact that the text would have contained no antecedent for the noun phrase "The real world."

At the other end of the scale is Old (what Prince calls Evoked) information. This is information I assume you already have in your understanding of our discourse (your discourse model, so to speak). I may assume that either because I've said it already, or because I assume you're aware of its presence in the immediate situation. In either case, I'm likely to use personal or deictic pronouns.

Finally, I may want to discuss information that is neither old--I haven't mentioned it before--or new--it's not completely unknown to you, and it's not in your head either. Instead, I assume that you can identify it by connecting it to some other entity in the discourse. This is what Prince terms Inferrable information. I may assume you can make an inference on purely logical grounds, but more likely I'll assume plausible reasoning. Inferrable information is always cast in definite noun phrases, but they may be either of two types. Containing Inferrables differ from Non-Containing Inferrable in that, in the former, the two entities to be connected occur within a single noun phrase. In the latter, they do not. Rather, I assume that you can make the appropriate connection without it being explicit.

In her analyses of written texts, Prince found that there was a good deal of "blurring between what is Unused and what is Inferrable" (1981, p. 251), because writers must usually take into account a group of readers knowledge and beliefs will differ from member to member. They need to give those readers with insufficient knowledge enough information to know what is being discussed; on the other hand, they don't want to give unnecessary information to more knowledgeable readers. Use of this strategy allows the writer to be co-operative with both kinds of readers. Because this strategy seemed so significant, I added it as a separate category in my study. Information, usually cast in the form of a complex noun phrase, that might be considered Unused for some readers but Inferrable for others was classified as Unused-or-Inferrable information.

I classified information represented by noun phrases in the essays according to their degree of assumed familiarity and also classified them according to their clause position--pre-subject, subject, or post-subject. Three entity types correlated significantly with essay score: Unused-or-Inferrable entities in all 3 positions (pre-subject: .33; subject: .34; post-subject: .33; all were significant at p= .001; the total number in all positions correlated at .42; p= .000); Containing Inferrables in subject position (.41; p= .000)[1]; and the total number of Brand New entities that were anchored syntactically to some other piece of information in the discourse varied less strongly (.26; p= .005), but still positively. Taken together, use of these three entity types accounts for almost 33% of the variance in essay score (r^2 =.325).

The significance of these results is twofold. They demonstrate, first, that readers' interpretations of the pragmatic function of structures in a text do influence their evaluation of its overall quality, i.e., if readers interpret writers as being co-operative, they reward that co-operation. But, further examination of the essays suggest that the situation itself influences how readers interpret writers: readers were not simply rewarding those writers who presented information most efficiently. Let me explain the first point.

In her analyses, Prince found "a preferred hierarchy or scale for what type of entity is used" (1980, p. 245), based on Grice's Maxim of Quantity. Writers give readers information as complete as necessary to identify an entity, assuming as much familiarity on the readers' part as

[1]See notes on page 162.

is reasonable. To be co-operative, writers will choose an entity type as high in the hierarchy as they can, as follows: (1) Old (evoked); (2) Unused; (3) Unused-or-Inferrable[2]; (4) Inferrable; (5) Containing Inferrable; (6) Brand-New. The appendix lists the full hierarchy and offers an example illustrating the principle.

Which one writers choose will depend on how familiar they believe readers are with the entity they are talking about. Using the example on page 159, if I've already discussed MacArthur earlier, I would most likely use (a). If I hadn't mentioned him but I assumed you knew of him, I'd use (b). If I assumed some readers may not have heard of him, I'd choose (c). And so on.[3] But, assuming that the purpose of our discussion is to give or receive information, I will try to give you no more or less information than I have reason to believe you need at that point in our discourse. In other words, the hierarchy reflects our attempts to be co-operative with each other, in the sense Grice (1975) uses.

The hierarchy developed by Prince predicts that writers who structure the information in their texts in a manner deemed cooperative by readers will be rewarded for doing so. And, in fact, the student writers I studied were evaluated positively to the extent they did so. The use of Containing Inferrables and Unused-or-Inferrable entities are sophisticated strategies for tailoring a message to the needs of unknown readers. However, the fact that readers rewarded the total amount of Brand-New, Anchored information and the kinds of Unused-or-Inferrable entities they rewarded raises questions about how Grice's Co-operative Principle is being applied in this situation.

Based on Prince's hierarchy, we might assume that writers were being rewarded for making clear and accurate distinctions between information completely unfamiliar to readers and information more familiar to them, while linking that information to topics previously introduced. But, that is not the case for these texts. Even the kind of information formulated as Brand-New differs in important ways from what Prince found in her own analyses. In the oral narratives that she analyzed, Brand-New, Anchored entities generally represented specific indefinites, references to some one individual or thing assumed unfamiliar to the reader, as in the sentence

John wants to marry a Norwegian, and there she is in the corner.

Those in the students' writing appear much more often to represent non-specific indefinites, if not generics. For instance, in the sample essay on page 160 in the appendix, of the 14 entities labeled Brand-New or Brand-New, Anchored, at least 9 represent non-specific indefinites. When this writer refers to "a group discussion," "a book report following questions written up by the school," or "a book from a school," she seems to be assuming that little of substance is shared with readers for an entire class of entities--an astonishing assumption given the topics of discussion in students' essays. Of course, the fact that students had had to write impromptu on topics unknown to them until they took the test made it difficult for them to fulfill the Maxim of Quantity, without at the same time failing to fulfill the Maxim of Quality, which states, in part, "Try to make your contribution one that is true" (Grice, p. 46). Their use of non-specific indefinites seems to be a way for writers to resolve the conflict, while the anchoring keeps up at least the appearance of keeping the discourse "coherent." That readers rewarded their use demonstrates that readers were evaluating, not only writers' communicative abilites, but also students' test-taking abilities--their ability to construct the form of communication even to the extent of sacrificing its substance.

In many instances, the use of Unused-or-Inferrable entities also seemed to represent more of a strategy for passing the test than for responding to the communicative demands of the audience for sufficient information. Consider the sample essay included on page 161 in the appendix.

This essay received a score of 12, the highest score possible on the test. But I did not choose this essay to argue that it should not have received this score; rather, I want to examine it as a clear example of a strategy by which writers can achieve such a score. In the essays I examined in this study, the kinds of information writers chose to portray as shared seemed to influence readers' evaluations a great deal. In this essay, for instance, on a supposed proposal to ban certain books from the town library, if the writer had wished to illustrate his point with examples of books typically the target of book-banners, he would likely have picked books by such popular, though explicit authors as Judy Blume. Instead, this writer refers to "Sophocles's 'Oedipus the King'" and "his (i.e., "Shakespeare") Othello." It is from these exemplars, and not from the likes of Judy Blume, that we are to infer the meaning of "fine literature." Similarly, at the end of the essay, the writer could have chosen many quotations, but the one he did choose only serves further to identify as shared that literate tradition from which the quote was taken. That is not a tradition likely to have been shared by many of the readers supposedly addressed by the topic, in this case citizens of a small town in rural Pennsylvania. Indeed, if they had been the actual readers, they might have interpreted these assumptions quite differently. But this writer knew well that the audience addressed was not the audience to be invoked.[4]

In the essay's opening paragraph, the writer introduces the term "censorship" to label the actions of those who propose to ban books from the school library. This term too seems to presuppose shared values as well as shared information. Certainly, those who propose to ban books would not accept it as applying to their actions.[5] Yet, readers' acceptance of that connection is crucial, because the term is central to the writer's elaboration of his position. If this writer had been less sure of the values he shared with readers, he presumably would have used a different term or would have "spelled out" the connection he was proposing between it and the school board's proposed action. But he did not, assuming instead readers who would be willing to accept that choice of terms without demanding that the writer defend, or at least explain, the basis of the connection. The score of 12, given to fewer than 2% of all the essays scored that summer demonstrates how accurate his assumptions were.

This use of Unused-or-Inferrable information as a way to appeal to shared values is, of course, important for any writer but would seem particularly important for writers who are students writing in a situation where they are also being examined. Cooper (1984, pp. 119-120), in discussing Grice's maxims, points out that the Maxim of Quality is both a sincerity condition--I have to establish that I believe what I'm saying--and a credibility condition--I have to establish that I have some authority for saying it. For some, the mere fact of their position establishes their authority. Clerics can marry, mayors can make proclamations, and teachers, I suppose, can lecture simply on the strength of their status in the situation. The status of these writers--students--gives them no such strength.[6] If anything, the opposite is the case. These writers are not presumed to be the kind of people who can explain/defend a position. They are assumed incompetent until they prove otherwise. And, if readers' evaluations are any evidence, the use of entities that appeal to shared traditions, especially literate traditions, seems to be a very successful way for writers to establish their credibility.

In fact, the need to establish credibility at times seems to drive writers to assume less familiarity with information than they have reason to believe is the case. Neither the date added to the phrase, "The Russian Revolution," nor the identification of the authors of Oedipus the King and of Othello, nor even the identification of Lord Acton as the author of the final quotation is necessary to identify the information. Readers who could not identify Oedipus or the source of the quotation are not likely to be able to identify the named authors. Few readers, on the other hand, do not know who wrote Othello. But, the appearance of this information does tell readers that the writer is the kind of student who knows these things. He can date the Russian Revolution and attribute authorship of great literature and literate quotations. In other words, when writers must choose in this situation between fulfilling the Maxim of Quantity (giving only enough information necessary for readers to identify information) or the Maxim of Quality (establishing their credibility), the second choice seems to be the one rewarded by readers.

The results of this study clearly show that readers do take into account the communicative function of text structures when they rate essays for overall quality. And, it also shows that they take into account the special demands of the testing situation. If the readers studied here are typical--and I have no reason to think they aren't--we can presume that they do not, generally, value papers that show "How to Say Nothing in 500 Words (Roberts, 1958). But, in this situation, such attempts are interpreted as appropriate to the demands of a test--a situation in which formal display takes precedence over substantive communication.

That readers even here value substance over form is evident in the extent to which they responded to the use of Unused-or-Inferrable information. But, here, too, the fact that this was a test seems to have influenced how they evaluated its use. In the example essay and in the others in the study, readers responded most positively to that information that identified writers as good students, who knew things and who shared the literate values of that community to which they were seeking entrance.

REFERENCES

Cooper, M. (1984). The pragmatics of form: How do writers discover what to do when? In R. Beach & L. Bridwell, New directions in composition research. (pp.109-126). New York: Guilford Press.

Diederich, P. B. (1974). Measuring growth in English. Urbana: National Council of Teachers of English.

Ede, L. & Lunsford, A. (1984). Audience addressed/audience invoked: The role of audience in composition theory and pedagogy. College composition and communications, 35, pp., 155-171.

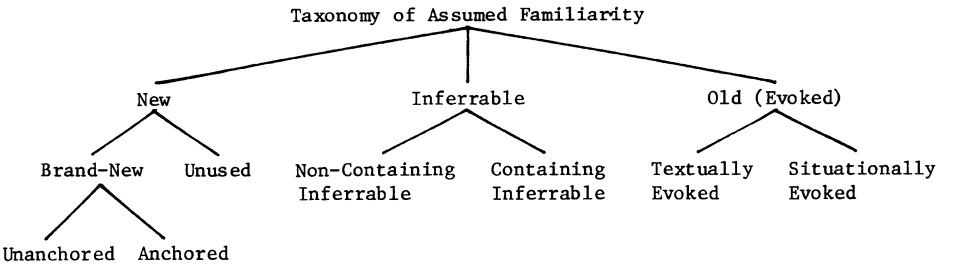
Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), Syntax and Semantics, 3, (pp. 41-58). New York: Academic Press.

Ohmann, R. (1982). "Reflections on class and language." College English, January, pp. 1-7.

Prince, E. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.), Radical Pragmatics (pp. 223-255). New York: Academic Press.

Roberts, P. (1958). Understanding English. New York: Holt, Rinehart, and Winston.

APPENDIX



A New: an entity first introduced into the discourse.

1. Brand-New: entity new in discourse and assumed unfamiliar to hearer.
   a. Unanchored: unconnected syntactically to any other entity.
      Example: A man went to Penn.

   b. Anchored: linked, through an NP contained within it, to another discourse entity.
      Example: A man that I know went to Penn.

2. Unused: entity new in discourse but assumed to be known to the hearer.
   Example: In the real world, discourse is judged by what it accomplishes.

B. Inferrable: entity assumed discoverable through logical or plausible reasoning from other entities in the discourse.

1. Non-Containing: the inference is to be made between entities in different syntactic constructions.
   Example: I walked into the classroom. The students were milling about.

2. Containing: the inference is to be made between entities contained within a single syntactic construction.
   Example: One of these eggs is rotten.

C. Old (Evoked): an entity assumed to be present in the current discourse model.

1. Textually: presence assumed because it has been introduced through the spoken or written text.

2. Situationally: presence assumed because it represents participants or feature of extratextual context, including text.

Adapted from E. Prince (1981), Toward a taxonomy of given-new information, in P. Cole (Ed.), Radical pragmatics. New York: Academic Press.

Hierarchy of Assumed Familiarity

Ranking of Entity Types in Order of Decreasing Familiarity Assumed by Their Use.

1. Old (Textually or Situationally Evoked)

2. Unused

3. Unused-or-Inferrable

4. Non-Containing Inferrable

5. Containing Inferrable

6. Brand-New, Anchored

7. Brand-New, Unanchored

To be co-operative (Grice, 1975) when referring to some entity, a speaker/writer "normally" uses an entity type as high on the scale as possible, depending on how familiar he/she assumes hearer/readers are with the information.

(a) He

(b) Douglas MacArthur

(c) Douglas MacArthur, the famous general came home

(d) The commander in disgrace

(e) The commander of U.N. forces in Korea

(f) A commander of U.N. forces in Korea

(g) A commander

Essay Relying on Brand-New, Anchored Information
(Total Essay Score of B)

Banning books due to unethical content has become a rising issue in high school libraries (BNA;S). The high schools don't want anything to do with books that contain foul language and/or explicit sex (BNA;NS). Instead of banning such books, the schools could use these books as a way of teaching the students what is actually contained in the books and how to deal with it (BNA;S). The students would read the books anyway, whether receiving them from school or somewhere else. Instead of having the students get the book from somewhere else and just reading the "trashy" parts, the schools could form some sort of program (BNA;NS). One idea could be a group discussion (BNA;NS). A few students (BN;NS) could read the book and then discuss it with a teacher, a librarian, etc. (BN;NS). Another program would be that if a student wanted to check the book out of a library he would have to get his parent's permission. The student would also have to write a book report following questions written up by the school (BNA;NS).

Sex, violence, and language have all become a big part in today's society (BNA;S). Sex, especially has become more outspoken, it is displayed on T.V., in magazines, in the movies, and in books. Banning these books aren't going to shelter the students over the issues. In my opinion it's just an easy way out for the high schools (BNA;S), one less problem to deal with.

I can see the school's point on one hand, that by keeping these books the parents of students might get upset. Even some of the parents are ignorant. They won't let their child read a book from a school (BNA;NS), but they will let them go to the movies where sex is displayed on the screen in front of the child's eyes.

The whole issue of banning books should be brought up before the school board, but the issue should be to keep the books; devise programs to teach the students what is in them, what the author was saying, etc. The issue should be talked over with the parents. Maybe the parents could read the books with their children and discuss what sex is about and what is ethical and what isn't.

Books should not be banned from high school libraries. If a student (BN) wants to read a book that isn't up to "standards," (BNA;NS) then there should be some instructionilized guidance to go along with it (BNA;NS), otherise, the student will get the book from somewhere else just to be rebellious and find out what it is that's so bad in the books.

Note: BNA = Brand-New, Anchored; BN = Brand-New, Unanchored; S = Specific; NS = Non-Specific.

Essay Relying on Unused-or-Inferrable Information
(Total Score of 12)

Topic: Suppose that your school board has proposed to ban certain books from the high school library on the grounds that they contain foul language or explicit sex. Write an essay for your local newspaper that explains to the school board your position as a student on this issue. Be sure to include good reasons for your stand.

The school board of Emmaus High School (UOI) has been considering a ban on certain books in its high school library. As a student of this school and a citizen of the free and democratic United States I must vehemently protest this action. If the school board votes to ban certain books because of lewd language or explicit sex, what is to stop them or other institutions from banning books for political, social or religious reasons? Limited censorship (UOI) can be a dangerous thing because it is a power that is very easily abused.

After the Russian Revolution of 1917 (UOI) one of the first things the new regime did was [to] ban books "offensive" to the government. Many of these books were not offensive because they contained explicit sex or foul language but because they conveyed ideas and principles that conflicted with those of the government. In many cases explicit sex and foul language were used as excuses for this censorship. In Nazi Germany (U) book burning UOI) was common. Today (UOI) Russia and other Communist-block countries (U) actively censor and burn books. Giving anyone the power to censor books is unwise, there is always the danger of getting carried away and grasping too much power, as the activities of Russia and other nations plainly show.

The censorship of books and repression of intellectual activity for whatever reasons are the first step toward a totalitarian regime. If school boards are given the right to censor books perhaps the government (UOI) will one day wish to ban books in schools and universities supported by federal funds. Once the wheel starts rolling it will be difficult to stop. The government might then wish to censor other aspects of citizens' lives.

Many works of fine literature contain foul language and explicit sex. Some of these works would be incomplete without them. Would one censor Sophocles's "Oedipus the King" (UOI) becuase it contains references to incest and also violence? Should Shakespeare (U) be banned because his "Othello" (UOI) portrays adultery (UOI)? Some people might say yes. What is not offensive to one may be shocking to others.

Most high school students have been exposed to foul language and explicit sex from a very early age. The electronic media (U) is greatly responsible for this. I believe that a few explicit paragraphs in a book will do little to enlighten an already worldly child to the evil ways of the world. If people are afraid [that] children will be shocked or offended by certain books then they can put little markers on them saying that these materials might be offensive to some people. These little markers might even work to some advantage. A child who rarely reads might be enticed to read some fine literature.

I believe that the school board should not be given the power to exercise censorship. Censorship in any form is an evil thing that can have catastrophic consequences. Lord Acton (U) put it best when he stated, "All power corrupts, but absolute power corrupts absolutely." (UOI) We must not allow this to happen.

Note; UOI = Unused-or-Inferrable; U = Unused.

NOTES

1. All statistics represent Pearson Product-Moment Correlations, unless stated otherwise.
2. Prince's analyses themselves did not distinguish this as a separate type of entity. I have extrapolated it from the blurring she observed between Unused and Inferrable entities. Its position in the scale follows from the Gricean Principle (1975) on which the scale is based.
3. Of course the situation is more complicated than that. Following Gricean notions, I may deliberately choose to use an entity lower or higher on the scale than my assumptions about the state of your knowledge would dictate. If I do so, then the burden falls on you to figure out the reason why I have not been fully cooperative. How such a choice affects readers' evaluation of the writing is discussed below.
4. The distinction between audience "addressed" and "invoked" is discussed in Ede and Lunsford (1984).
5. It is true that they would have understood that the writer had intended it to so apply--the illocutionary act it was supposed to perform. Yet, it is not clear that they would have accepted the writer's right to perform it. There doesn't seem much difference between someone saying, in this case, "You can't tell me that. Who do you think you are anyway--a know-it-all?" and someone saying, "You can't marry us! Who do you think you are, anyway--a preacher?"
6. This is true, of course, not just of placement-testing in particular, but of any situation in which there are marked asymmetries in power between the participants. See Richard Ohmann (1982) for an insightful analysis of the effects of such asymmetries on language use.

A COMPARISON OF BAYESIAN AND TRADITIONAL INDICES FOR MEASURING INFORMATION GAIN SENSITIVITY IN A CLOZE TEST

Kyle Perkins and Worthen N. Hunsaker

Southern Illinois University
Carbondale, IL

Standardized reading comprehension tests have been criticized for their inability to assess what a reader has gleaned from the passage during the reading process (Tuinman, 1973). Tuinman reported scores as high as 65 percent correct on multiple-choice comprehension tests which subjects answered without having had access to the passages on which the questions were based. Tuinman's paper, among others, has caused reading researchers to focus on the distinction between comprehension and information gain (IG). In general, the concept of IG involves comparing the reader's state of knowledge before and after reading the test passage. One method of assessing IG, suggested by Coleman and Miller (1967), is to obtain percent correct cloze scores before and after the subjects have had an opportunity to read the test passage in its original, unmutilated form; the difference between the posttest scores and the pretest scores is considered to be the IG score.

The purpose of this study was to apply Bayes' Theorem to item analysis of cloze items and to compare the resulting indices of item effectiveness with criterion-referenced and classical test theory indices. The focus of the study is the performance of these indices for measuring IG sensitivity. The indices are: (1) the pre-to-post difference index (PPDI) introduced by Cox and Vargas (1966), (2) the percent of possible gain (PPG) introduced by Brennan and Stolurow (1971), (3) three indices based on the Bayes' Theorem presented in Haladyna and Roid (1981), (4) the Brennan index, or BI (Brennan, 1972), (5) three test item discrimination indices, D1, D2, and D3, suggested by Helmstadter (1974), and (6) four indices, 01, 11, 00, and 10, introduced by Popham (1971). These indices have previously been employed to assess "the tendency for an item to vary in difficulty as a function of instruction" (Haladyna and Roid, 1981: 40). The present study was planned to utilize these indices which were calculated on cloze test items and to examine the role played by IG in affecting subjects' performance.

PPDI is the difference in item difficulty (determined as the proportion of correct responses) observed when an item was given to subjects before they had read the unmutilated passage, and then after they had read the passage in its original form. According to Roid and Haladyna (1982: 218-219), "this index is the simple difference between the difficulty observed in the post-instruction group and the difficulty observed in the pre-instruction group. It therefore ranges from -1.00 to +1.00, with indexes

of zero or lower quite rare. In most instructional settings, this index typically ranges from .10 to .60."

PPG is a variation of PPDI, and it is computed by the formula

$$PPG = \frac{\text{difficulty (posttest)} - \text{difficulty (pretest)}}{1 - \text{difficulty (pretest)}}$$

PPG "is based on the notion that with any pretest item difficulty, there can only be a maximum gain, which is 1.00 minus the pretest difficulty" (Roid and Haladyna, 1982: 219).

Helmstadter (1974) introduced three discrimination indices based on the Bayes' Theorem (Bayes, 1763). Bayes, an English clergyman who died in 1760, produced a formal derivation of the equality among certain probabilities, and the theorem "is based on the fact that the joint probability of two events P and D can be written as the probability of one of the events and conditional probability of the second event, given the first event" (Iversen, 1984: 12).

The aforementioned prose can be expressed in the following equation:

$$\text{Prob (PD)} = \text{Prob (P)} \cdot \text{Prob (D|P)}.$$

The joint probability of the two events can be paraphrased as follows if the two events are reversed.

$$\text{Prob (DP)} = \text{Prob (D)} \cdot \text{Prob (P|D)}$$

The two left sides are equal, and the two right sides are equal. By equating the right sides and by rearrangement, the following equation is produced:

$$\text{Prob (P|D)} = \frac{\text{Prob (D|P)} \cdot \text{Prob (P)}}{\text{Prob (D)}}$$

There can be several events P, i.e., P1 through Pk, which are both exclusive and exhaustive. The probability of D in the denominator is a weighted sum of the conditional probabilities Prob (D|Pi) where the weights are Prob (Pi).

The Bayes' Theorem for k different P's is:

$$\text{Prob (Pi|D)} = \frac{\text{Prob (D|Pi) Prob (Pi)}}{\text{Prob (D|P1) Prob (P1)} + \ldots. + \text{Prob (D|Pk) Prob (Pk)}}$$

The three indices based on the Bayes' Theorem requires collateral information; the most appropriate type of information is preinstruction test results. In the context of the present study, the pretest item difficulty is drawn from a sample of subjects who have not read the test passage in its unmutilated, original form. The three Bayesian indices are: (1) B1, the probability that a subject knows the content material given that the correct response was selected, (2) B2, the probability that a subject does not know the content material given that the incorrect response was selected, and (3) B3, the probability that a correct decision will be made about the subject's knowledge of the content given the results of performance on that item, that is, mastery or nonmastery. An implication is that items with high B3 indices augment the test administrator's probability of correctly classifying mastery and nonmastery subjects.

Helmstadter (personal communication) has supplied the following procedures for the development of the indices based on the Bayes' Theorem.

Definitions

R = gets item correct

$\bar{R}$ = gets item incorrect

K = has knowledge

$\bar{K}$ = does not have knowledge

Possible Outcomes

	P(K)	P($\bar{K}$)
	Knows Material	Doesn't Know Material
P(R) gets item correct	correct positive	false positive
P($\bar{R}$) gets item incorrect	false negative	correct negative

Bayes' Theorem for decisions for items based upon item performance

$$P(K|R) = \frac{P(R|K)\ P(K)}{P(R|K)\ P(K) + P(R|\bar{K})\ P(\bar{K})}$$

$$P(\bar{K}|\bar{R}) = \frac{P(\bar{R}|\bar{K})\ P(\bar{K})}{P(\bar{R}|\bar{K})\ P(\bar{K}) + P(\bar{R}|K)\ P(K)}$$

P (Correct decision) = P($\bar{K}$ and $\bar{R}$) or P(K and R)

$$P(RK) = P(R|K)\ P(K)$$

$$P(\bar{R}\bar{K}) = P(\bar{R}|\bar{K})\ P(\bar{K})$$

Haladyna and Roid (1981: 43) provide the computing formulae which entail knowledge of pretest and posttest item difficulties:

$$B1 = \frac{(POSTDIFF)\ (COMDIFF)}{(POSTDIFF)\ (COMDIFF) + (PREDIFF)\ (1\text{-}COMDIFF)}$$

$$B2 = \frac{(1\text{-}PREDIFF)\ (1\text{-}COMDIFF)}{(1\text{-}PREDIFF)\ (1\text{-}COMDIFF) + (1\text{-}POSTDIFF)\ (COMDIFF)}$$

$$B3 = (POSTDIFF - COMDIFF) + (1\text{-}PREDIFF) + (COMDIFF - POSTDIFF)$$

where

B1, B2, and B3 are the three Bayesian indexes, and
PREDIFF is the preinstruction sample difficulty,
POSTDIFF is the postinstruction sample difficulty, and
COMDIFF is the mean of PREDIFF + POSTDIFF, or the
combined samples' difficulty.

Haladyna and Roid claim that the conceptual basis for each index lies in the calculus of conditional probabilities.

The Brennan index, or BI has its basis in the dichotomy of a mastery-nonmastery contrast. A criterion level is established on the test scale, and all subjects having scores above this benchmark are classified as masters; all subjects having scores below this criterion are classified as nonmasters. The formula for the Brennan index is:

BI = difficulty (mastery) - difficulty (nonmastery)

D1, D2, and D3 are three separate variants of a classical item discrimination index which is computed using the sample separation technique. These three variants involve defining "high knowledge" and "low knowledge" groups in different ways. The following descriptions come from Helmstadter (1974).

D1 assumed that persons in the top one-third of the class on the posttest were "high knowledge" and persons in the bottom one-third of the class on the posttest were "low knowledge."

D2 involved combining scores from the pre- and posttest as if they constituted one large class and then assuming that persons in the top one-third of this doubled class were "high knowledge" and that persons in the bottom one-third of this doubled class were "low knowledge."

D3 assumed that the pretest scores represented a "low knowledge" group and that posttest scores represented a "high knowledge" group.

The 01, 11, 00, and 10 indices were introduced by Popham (1971). When subjects take a pretest and a posttest, there are four logical states of affairs: (1) 01, an incorrect response on a pretest and a correct response on the posttest, (2) 11, a correct response on the pretest and a correct response on the posttest, (3) 00, incorrect response on both tests, and (4) 10, a correct response on the pretest and an incorrect response on the posttest. The percentage in the 01 cell for subjects who took both pretest and posttest would provide an index of instructional sensitivity.

Roid and Haladyna (1982: 220) provide a fourfold table to represent the four possible patterns of item responses from a pretest and a posttest.

	Posttest	
Pretest	Correct	Incorrect
Incorrect	01	00
Correct	11	10

METHOD

Subjects

This study was conducted in the Department of Linguistics at Southern Illinois University. Forty-one full-time undergraduate foreign students served as unpaid volunteers for this study. The subjects were freshmen and sophmores. Their native languages included Arabic, Brunei, Chinese, Greek, Korean, Malay, Spanish, Swedish, Thai, Urdu, and Yoruba.

To be admitted for full-time undergraduate study at Southern Illinois University, a foreign student must meet one of the following criteria: (1) have a Test of English as a Foreign Language score of 525 or higher, (2) pass the proficiency examination (a score of 70 or higher on the Michigan Test of English Language Proficiency and Aural Comprehension Test) and reading and writing examinations of the Center for English as a Second Language at Southern Illinois University, or (3) have completed 60 semester hours of collegiate training in an accredited United States college or university.

Material

A thirty-five item cloze test was constructed for this research by deleting every eighth word beginning with the thirty-second word in the passage. The passage was 348 words long, and the topic was Finland's form of government. The two first and two last sentences were left unmutilated.

The distribution of deletions was the following: 9 nouns, 11 verbs, 7 adjectives, 3 articles, 1 preposition, 1 conjunction, and 3 pronouns.

Procedure

As a pretest, the cloze was administered in class. Working time was forty minutes. Three weeks later, the subjects were given an unmutilated version of the entire passage to read and study for twenty minutes. The passages were collected, and a fresh copy of the same cloze test was administered a second time; forty minutes was allotted for the test. The reading and posttesting were conducted in class. Exact-word scoring was used to mark both the pre- and posttests. No indication was given during the pretest that the subjects would be tested again.

Cloze researchers recommend that a cloze test have at least 50 items; however, the pilot testing for this project indicated that 35 items was about the threshold at which response arbitrariness began to manifest itself for foreign students similar to those who comprised the sample for this study.

In the context of this project, the subjects' reading and studying the unmutilated passage is equated with instruction. The pretest is equated with preinstruction, and the posttest, with postinstruction.

The value for each of the thirteen indices was calculated for each of the thirty-five cloze deletions according to the description for each index. Only one index represents a caveat, the Brennan index. We chose 80 percent correct as the criterion for mastery, that is to be declared a master a subject must have a score of at least 28 (35 items x .8 = 28). For the pretest there were no masters. For the posttest there were twelve masters and twenty-nine nonmasters; therefore, the Brennan index for this study was based only on the posttest.

RESULTS AND DISCUSSION

Table 1 contains the means, standard deviations, standard errors of measurement, and estimates of reliability for the pretest and posttest person scores. The posttest scores were significantly higher than the pretest scores, $t(40) = 9.64$, $p < .0005$, one-tailed test. The observed estimates of internal consistency are not pleasing, but the estimate of reliability for the posttest (.80) is higher than the pretest estimate (.52). This finding runs counter to the normal expectation because gain scores entail using a test twice, and therefore, there are two opportunities for measurement error to creep in. It is highly probable that the subjects' reading the unmutilated passage and their having become familiar with the cloze procedure better enabled them to utilize the inherent redundancy in the text, thereby enabling them to analyze by synthesizing better the second time around. The result is that the posttest data more truly exhibits the diversity of the subjects' processing ability, and consequently, there was better person separability or reliability.

Table 2 presents the medians and ranges for each of the indices. Table 3 portrays the intercorrelations of the various indices of item effectiveness. Four of the indices are direct measures of IG: PPDI, PPG, D3, and 01. PPDI had significant correlations with PPG, B1, B2, B3, D3, 01, 11 (negative), and 10 (negative). PPDI is a direct measure of IG, is simple to compute, and is consistently and highly related to most of the other indices.

PPG correlated significantly with B1, D3, 01, and 10 (negative); however, PPG is suspect because it is distorted when the pretest difficulty is high.

D3 correlated significantly with 01, 00, 11 (negative), and 10 (negative). 01 is not nearly as robust as PPDI or PPG. In general, there was a lower degree of relatedness between the traditional indices and the other indices.

We noted the following generalizations about the performance of the Bayesian indices. They are highly influenced by item difficulty. B1 is high when pretest and posttest sample difficulty is high. Pretest sample difficulty affects B2 to the extent that when pretest difficulty is low, B2 is high. When the combined-samples difficulty is high, B2 is low. And finally, B3 is fairly reflective of consistent gains; in fact, B3 had the highest mean correlation in terms of magnitude (ignoring the sign of the coefficient) of all the indices. The fact that the Bayesian indices are influenced by item difficulty and are unstable across samples, according to Haladyna and Roid (1981), make them suspect for all but research purposes.

Table 1. Means, Standard Deviations, Standard Errors of Measurement, and Estimates of Reliability

	Number of Subjects	Number of Items	$\bar{X}$	s	s_e	KR-20 r_{tt}
Pretest	41	35	12.39	3.18	2.20	.52
Posttest	41	35	22.51	5.55	2.48	.80

Table 2. The Medians and Ranges of the Various Indices

Index	Median	Range
PPDI	.24	.00 - .88
PPG	.40	.00 - 1.00
B1	.77	.35 - 1.00
B2	.66	.00 - .92
B3	.68	.05 - 1.00
BI	.23	-.04 - .59
D1	.61	.45 - .92
D2	.61	.46 - .92
D3	.63	.50 - 1.00
01	.27	.05 - .88
00	.34	.00 - .68
11	.25	.00 - .95
10	.02	.00 - .15

There are other facets of the Bayesian indices that vex us greatly. Haladyna and Roid define B1 as the probability that the subject has knowledge, given that the subject answers the item correctly. It follows from the nature of the cloze test and exact-word scoring that B1 should equal 1 for each subject in our sample who answered correctly. We note that for our sample the right hand side of the B1 formula lies in the interval of .35 to 1.00, and therefore, we question the validity of this estimate of B1.

It also follows from the nature of the cloze procedure, exact-word scoring, and the Bayes' Theorem that B2, the probability that the subject does not have knowledge, given that the subject misses the item is also equal to 1. The data from this project indicate a range from .00 to .92 for the right hand side of the formula to estimate B2 given by Haladyna and Roid. And finally, B3 is not Bayesian; it is simply the unconditional probability that the subject has prior knowledge.

DIRECTIONS FOR FUTURE RESEARCH

We intend to continue our research with these indices along the following lines. There is no formal statistical decision theory to indicate what the acceptable ranges for each of these indices should be. We know, for example, that the acceptable range for item difficulty is .33 to .67, that the minimum discrimination index for the sample separation procedure is .67, and that the minimum point-biserial correlation for item discriminability is .25, but we do not have similar criteria for the indices studied in this project. The purpose of instructional sensitivity study is to assist the researcher in determining whether the instruction has been faulty or whether an item is flawed or inappropriate. Guidelines are needed for these indices before serious work can be attempted on instructional sensitivity.

We would also like to replicate this study with another cloze test (hopefully a more reliable one) and apply a principal axis factor analysis with varimax rotation to each intercorrelation matrix, thus obtaining an independent factor structure for each of the two samples.

Table 3. Intercorrelations* Among Various Indices of Item Effectiveness

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. PPDI		.657	.439	.764	.774	.152	.073	.136	.863	.985	.178	-.717	-.667
2. PPG			.578	.234	.235	-.176	-.163	-.157	.446	.649	-.299	-.149	-.553
3. B1				-.051	.020	-.051	-.217	-.167	.439	.397	-.371	.059	-.421
4. B2					.966	.436	.385	.495	.834	.769	.687	-.968	-.481
5. B3						.478	.415	.529	.898	.758	.753	-.989	-.601
6. BI							.842	.857	.431	.124	.602	-.478	-.308
7. D1								.852	.321	.046	.588	-.422	-.257
8. D2									.438	.096	.714	-.533	-.355
9. D3										.821	.535	-.848	-.756
10. 01											.144	-.724	-.538
11. 00												-.776	-.366
12. 11													.485
13. 10													

*For 33 df the 1% critical value of r is .409.

It is well known that cloze item responses are inversely related to uncertainty in a passage. Our future research will also investigate the sensitivity of these indices for measuring the IG of word classes which transmit the most new information in relation to cloze item difficulty attributable to three characteristics of the deleted word: (1) common logarithm of the frequency of a "word family's" appearance in large samples of written material, (2) the degree to which the word is likely to be encountered across different content areas, e.g., the humanities, social sciences, science, etc., and (3) whether the word is abstract or concrete.

REFERENCES

Bayes, T. 1763. Essay towards solving a problem in the doctrine of chances. Philosophical Transactions 53, 370-418. (Reprinted in Biometrika, 1958, 45, 293-315.)

Brennan, R. L. 1972. A generalized upper-lower item-discrimination index. Educational and Psychological Measurement 32, 289-303.

Brennan, R. L. and L. M. Stolurow. 1971. An empirical decision process for formative evaluation. Research Memorandum No. 4. Cambridge, MA: Harvard University CAI Laboratory.

Coleman, E. B. and G. R. Miller. 1967. A measure of information gained during prose learning. Reading Research Quarterly 3, 369-386.

Cox, R. C. and J. Vargas. 1966. A comparison of item selection techniques for norm-referenced and criterian-referenced tests. Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago.

Haladyna, T. and G. Roid. 1981. The role of instructional sensitivity in the empirical review of criterion-referenced items. Journal of Educational Measurement 18, 39-52.

Helmstadter, G. C. 1974. A comparison of Bayesian and traditional indexes of test item effectiveness. Paper presented at the annual meeting of the National Council of Measurement in Education, Chicago.

Iversen, G. R. 1984. Bayesian statistical inference. (Sage University paper No. 43. Series; Quantitative Applications in the Social Sciences). Beverly Hills, CA: Sage Publications, Inc.

Popham, W. J. 1971. Indices of adequacy for criterion-referenced tests. In Popham, W. J., editor, Criterion-referenced measurement. Englewood Cliffs, NJ: Educational Technology Publications.

Roid, G. and T. Haladyna. 1982. A technology for test-item writing. New York: Academic Press.

Tuinman, J. J. 1973. Determining the passage difficulty of comprehension questions in 5 major tests. Reading Research Quarterly 9, 206-223.

THE C-TEST: A VIABLE ALTERNATIVE TO THE USE OF THE CLOZE PROCEDURE IN TESTING?

Mary Ann G. Hood

English Language Institute
The American University
Washington, D.C.

A few years ago, two German researchers, Christine Klein-Braley and Ulrich Raatz, developed a test procedure which they believe maintains the virtues of the cloze test procedure while exhibiting none of its defects -- the C-Test. The C-Test consists of a number of texts or passages in which, beginning in the second or third sentence in each passage, the second half of every other word has been deleted, leaving the first part of each mutilated word as a cue for the test taker.

Klein-Braley and Raatz set out for themslves a set of specifications for the development of their test, including the following:

a) It should be a test of reduced redundancy.
b) It should make use of "authentic" texts.
c) In such a test, it should be possible to make use of shorter texts or passages, but each test would have at least 100 items.
d) The deletion rate and the point at which deletion begins should be fixed.
e) The words affected by deletion should represent a genuine sample of all the elements in the text.
f) Only exact scoring should be permitted so as to ensure maximum objectivity.
g) Various types of texts should be used so as to assure that no individual test taker has a particular advantage.
h) Native speakers ought to be able to achieve perfect or nearly perfect scores on such a test.

Much of the investigation of the usefulness of the C-Test has been carried on in Germany, in native language testing projects and second and foreign language testing projects, involving elementary school children, secondary school students, and adults. C-Test projects have also been tried in a few other European countries and in Israel, involving, to date, a total of six different languages. As shown in a fairly recent article by Klein-Braley and Raatz, the results are impressive. (See Christine Klein-Braley and Raatz, "A Survey of Research on the C-Test," Language Testing 1, no. 2 (Dec. 1984): 134-146.)

Based on such C-Test construction guidelines as I could derive from the published work of Klein-Braley and Raatz which was available to me,

I set about to construct a modest C-Test and Cloze Test experiment. Specifically, the C-Test and Cloze Test passages in this project were chosen according to the following guidelines:

1. Both tests would have the same number of passages, five.

2. Test passages for both tests would involve similar types of content; thus, both tests have passages whose content deals with health, economics, marketing, mathematics, and computer science. At my institution, economics, marketing, mathematics, and computer science are popular major fields of study for foreign students. Health was chosen for the fifth content area as being a kind of 'neutral' subject.

3. Test passages were selected from college or university level materials; that is, the text sources are of the type that can be found in the library at my university. The test passages represent a variety of reading levels, from 9th-10th grade level to graduate level.

4. With one minor exception, none of the test passages was modified in any way. (The minor modification involved interpreting a dash in the original text as a sentence marker; that is, in the test passage, the dash became a period and what follows constitutes a complete sentence, which it did, grammatically, in the original material.)

5. The total number of deleted or mutilated items in both tests involves nearly the same proportion of structure words and content words (about half and half in each test).

Obviously, these guidelines suggest that the test I wished to construct should not be administered to students at the lower end of the English proficiency continuum, and that is precisely what I intended. These tests were meant only for those non-native English speaking students whose command of English is said to be at the high intermediate or advanced level of proficiency.

What I hoped to determine, at least up to a point, was the following:

Does the major field of study of the non-native English speaking student give that student any particular advantage in dealing with certain portions of either test?

Conversely, is the subject matter of the test passage of little significance if the test-taker is a "good reader" and/or very proficient in English?

Does the test-taker's native language appear to have any significant impact on his/her score on the C-Test?

Would educated adult native speakers achieve perfect or virtually perfect scores on the C-Test passages?

How significant a factor would misspelled words turn out to be on the C-Test?

Before moving on to the results of this testing project, I should like to comment on some of the problems that I encountered--problems that seem not to have occurred in the testing projects reported by Klein-Braley and Raatz.

The first is the matter of the level of reading difficulty in the texts chosen for the C-Test. In developing their test, Klein-Braley and

Raatz used passages that might be said to represent, in a very general way, an elementary reading level (in German and English), an intermediate level, and an advanced level; this is clear from their reported research in 1984 and 1985. (Klein-Braley and Raatz, 1984; Christine Klein-Braley, "A Cloze-up on the C-Test," Language Testing, 2, no. 1 (June 1985): 76-104.) Further, the 1985 article by Klein-Braley includes rather extensive comments on sentence length--acknowledged by Klein-Braley to be a reasonable predictor of C-Test difficulty--and the type-token ration (which relates to vocabulary range). On the other hand, so far as I have been able to determine, the C-Test texts used by Klein-Braley and Raatz have not been ranked as to general level of difficulty or grade level. As I am sure some ESL reading specialists would point out, the grade or difficulty level may not be a critical issue--indeed, such ranking may not be appropriate for non-native speakers of English because it is not at all clear that such rankings mean the same for them as for native speakers. But the fact remains that what ESL students are ultimately required to do, if they wish to pursue undergraduate or graduate degree work in the United States, is to be able to read college or university level materials.

For this reason, my choice of texts for a C-Test was quite deliberate: the texts were to be chosen from "authentic" college level texts. As mentioned earlier, these texts range in level of difficulty from the 9th or 10th grade to postgraduate level, as determined by the Dale/Chall reading formula. (For those of you unfamiliar with the Dale/Chall formula, it does depend, in large part, on sentence length and the number of unfamiliar words; the formula does regard the most common function or structure words as familiar.) Given the difficulty of the C-Test passages, and the Cloze passages, it is not surprising that a number of foreign students who took the test did not achieve very high scores; on the other hand, some of them did make fairly high scores: overall, for the mixed foreign student group, percentage scores on the C-Test ranged from 33.% to 92.%, and for the Chinese group, percentage scores ranged from 31% to 87%.

However, the level of difficulty may also have affected the scores of the native speakers who took the C-Test. A total of 25 native speakers of English, all adults, tried the C-Test; some of them are undergraduate students (sophomore level and above), and about one-third are professionals, including some ESL teachers and lawyers. In other words, this group seems to meet the specification by Klein-Braley and Raatz that the native speakers who take a C-Test should be educated adults, thereby assuring perfect or nearly perfect scores. The educated adults who took my C-Test achieved percentage scores ranging from 75% to 97.7%. I would agree that scores above 90% can probably be viewed as "nearly perfect," but I find it hard to include in the nearly perfect range those scores between 75% and 89%.

A contributing factor to the less than perfect scores among the native speakers may have been spelling errors. Only two of the 25 native speakers had no spelling errors--one professional and one student. The other 23 native speakers had from one to five spelling errors, though most had only one such error. Spelling was also a problem for some of the non-native speakers who took the C-Test; a number of them had four or five or even more spelling errors, while many spelled one or two words incorrectly.

I have concluded, perhaps incorrectly, that Klein-Braley and Raatz would say that misspelled words in a C-Test must be counted wrong. In 1981, they reported on a German C-Test administered to German school children in the 3rd grade; among other findings, they stated that the seven dyslexic children who took the C-Test scored significantly lower than non-dyslexics--the implication is that poor spelling was the reason; and they also mentioned that one measure for identifying dyslexics

in Germany is an orthography test. There is, in this report, at least the implication that incorrect spellings constitute incorrect responses. (Ulrich Raatz and Christine Klein-Braley, "The C-Test: A Modification of the Cloze Procedure," Occasional Papers: Practice and problems in Language Testing no. 26 (1981): 113-138.) More recently, Raatz stated that: "C-Tests can be used as global diagnostic instruments of learning difficulties in language, both in the native language (dyslexia) and in the foreign language...." (Ulrich Raatz, "Tests of Reduced Redundancy--The C-Test, A Practical Example," Language Testing in School: AF in LA Yearbook 1985 (Tampere 1985): 49-62). Thus, given that Klein-Braley and Raatz specify that C-Test scoring must be exact, I have assumed that misspelled words must be counted as incorrect.

Based on my experience with this test project, such a rule, with respect to spelling, makes sense if the test-taker keeps constantly in mind that he/she has on the page a cue consisting of precisely one half a word, or that half the word plus one letter is missing. But for poor spellers (albeit reasonably good readers and proficient users of the language), this can be a problem. For example, in passage #2 of the C-Test, the passage whose subject is health, we find this sentence: "The amount of hemoglobin in the blood varies slightly between people, particularly between the sexes." The deletion pattern required that the word particularly be mutilated. Thus, the cue is partic_____ (6 letters in length), with six letters (__ularly) omitted. A number of foreign student test-takers--perhaps including some who conscientiously applied the rule--completed the word with --ularlly. This response was counted wrong, but I cannot help but wonder if it should have been counted as correct.

With respect to the native speakers who took the C-Test, it is difficult to know why some words were misspelled; my guess is that the errors occurred either through sheer carelessness or inattention. Frequent types of errors involved words which are often misspelled: resistence with an -e instead of resistance; flaging with one -g rather than two; arithmatic with an -a instead of arithmetic, etc. And even among the native speakers, including some of the professionals, the word particularly appeared in several variations.

Another problem was presented by the possessive form of one of the mutilated words, the word chip's in the passage devoted to computer science (passage no. 3). In the text examples presented by Klein-Braley and Raatz, I could find no example of a mutilated possessive noun; thus, I was not sure how to account for such a form in the directions for the C-Test, and I chose to ignore the matter. A number of native speakers were able to complete the word correctly, based on the cue given: chi_____. In other words, the cue which was presented treated the apostrophe as a letter. For this reason, this item was included in the calculations for tests completed by native speakers of English, but not for non-native speakers: none was able to complete the word correctly, not even the computer science majors.

From my point of view, this small matter does raise an interesting question because the possessive noun form is not an uncommon occurrence in English. How, then, should one deal with it in directions for the C-Test? Not to make mention of the possible need for an apostrophe, as I did not, seems to undermine the clarity of the directions in such a test. In fact, a few native speakers complained on this account. But, to include in the directions the possibility that apostrophes may be needed and that they are included in counting the 'letters' to be deleted might create more confusion than clarity for non-native speakers of English. Another possibility, of course, is to choose only those tests in which no possessive noun occurs, but this would seem to undermine the specification that

authentic texts can be used as test passages in a C-Test.

Another common type of error among both non-native and native speakers of English had to do with singular and plural forms. For example, in the first passage of the C-Test, the test taker is presented with the cue la____; the correct response is land, but several test takers used the plural lands. Such a completion does conform to the test directions, and such a completion is not grammatically incorrect in the particular context, though stylistically, it may leave a bit to be desired. There were similar problems with registers--the plural, not the singular is correct, though either would appropriately fit in the particular context. On the other hand, a careful reading of the entire passage ought to have suggested to the test taker that the plural form was the correct choice.

Still another type of error committed by both native and non-native speakers was to complete an item in a way that is both semantically and grammatically correct, and in conformity with the directions, but given the original language of the text, the responses were counted as incorrect. Examples include incorrect responses of teach for tell--the cue was te____; of for on--the cue was o_____; and those for these--the cue was th_____.

Klein-Braley and Raatz do deal with some of these problems, but not as fully as they might have, it seems to me. In a 1981 article which describes their progress in the development of the C-Test, Raatz and Klein-Braley stated that "The scoring was to be exact; in any case, because of the half-words that are left standing, there are very few cases in which alternatives are possible." (Raatz and Klein-Braley, 1981, p. 124.) Later on, in the same article, in discussing a pilot C-Test administered in Leeds, England, they reported that: "For 7 of the items on the test, there are viable alternatives. In particular, the deictic elements this/that permit alternative solutions. However, in no case is more than one alternative possible..." (Raatz and Klein-Braley, 1981, p. 125). Of a German C-Test, they stated: "For 98 of the blanks, there were no alternative solutions. In two blanks, there were two possibilities: both the singular and the plural of the mutilated word. Consequently, we decided that, in the future, both solutions would be counted as correct..." (Ibid., p. 129).

In more recent articles, Klein-Braley and Raatz continue to insist that only "exact scoring should be possible" (LT, 1984), and "Only exact scoring should be possible so as to ensure objectivity" (Raatz, 1985, p. 50), but also, "C-Test scoring is 'exact' and objective because there is almost always only one possible solution. In a few exceptional cases, there are a small number of different solutions (e.g., this and that).In such cases, all solutions are accepted." (Ibid., pp. 50-51.)

A final problem from my point of view is the extent to which most of the test takers seemed quite careless about completing the mutilated words: nearly every test exhibited instances in which the pattern for deletion, and the consequent rule for restoration, was ignored; i.e., the directions for the C-Test specify that, "Beginning in the second or third sentence in each passage, half of the letters (or half plus one letter) of every other word have been deleted. For example, a six letter word will have the last three letters deleted, and a seven letter word will have the last four letters deleted." This phenomenon--ignoring the rule for restoration of mutilated words--is mentioned only in passing by Klein-Braley (LT, 1985, p. 99) and in connection with a group of 10 year old English school children; in the same article, however, she does report that an investigation has begun on language processing strategies vis-a-vis the C-Test. It may be that the results of that investigation will clarify some of the matters which have puzzled me. In many instances, especially among native speakers, mutilated words were completed in such

a way as to present a real word, but generally a word that did not fit the context, or, if it did fit the context, it clearly violated the deletion pattern. In fact, two of the native speakers who took the test told me that they paid little or no attention to the number of letters in the cue; as they put it, they simply responded "by instinct."

Perhaps this mode of response--to fill in the blanks by instinct--is of little significance, but I do wonder if native English speaking Americans--at least my small sample--responded to the C-Test in a less than methodical way than other nationality groups.

I should like, now, to turn to the results of this modest testing project. In both tests, as I mentioned before, there were five passages with content drawn from economics, marketing, mathematics, computer science, and health texts. The C-Test had a total of 180 mutilations, but the final score, for non-native speakers, was based on 179 mutilations. Of the 179 mutilations or items, 90 involved content words, and 89 involved structure words; it should be noted, however, that the distribution of content words and structure words in the individual passages was not as even as these figures suggest.

For the Cloze test, 75 words were deleted: 36 content words (48%) and 39 structure words (52%). In general, the distribution of content words and structure words throughout the individual passages was somewhat more even than in the C-Test passages.

All types of words (nouns, verbs, adjectives, adverbs, etc.) were designated for deletion or mutilation in the two tests, though the C-Test, as might be expected, had a more even and broader distribution of token-types, simply because of the number of items.

The tests were administered to 58 non-native speakers of English, all but one of them being students at The American University and Georgetown University. Currently, these students are taking English either in high intermediate or advanced level classes; those in the advanced level classes are all taking classes in other subject areas, as undergraduate or graduate students. The 58 non-native speakers represent 19 different native languages, with Spanish, Arabic, and French being the most common languages, in that order. The students designated 12 different disciplines as major fields of study, the most common being business administration, computer science, economics, finance, and agriculture/agronomy.

The same two tests were also given to a group of Chinese English teachers, participants in a summer TEFL program at Hunan Medical College in Changsha, China; 28 in this group completed both tests. In their undergraduate studies, all had majored in English or were still studying English, and all are currently teaching English as a foreign language.

The test takers were also asked to complete a brief questionnaire regarding reading habits in both their native language and in English. The results of the questionnaire were, in general, not very useful, perhaps because the questions were not framed as well as they should have been.

For several reasons, I have chosen to deal separately with the test results obtained from those who are foreign students in the Washington D.C. area, and from the Chinese English teachers. While all in the Chinese group are English teachers, none has ever been outside China, and obviously, none has ever lived in an English speaking community. Further, I believe it must be conceded that foreign students who attend universities in the Washington D.C. area probably represent quite advantaged social and economic levels in their own countries, including the advantage of frequent

contact with other countries and cultures, which is not the case with the Chinese group. Thus, the two groups must be perceived as distinct test populations.

With respect to the question of whether the student's native language might have any significance with respect to C-Test and Cloze test scores, in the foreign student group, the answer is probably no, though I may not have gathered enough information to truly answer this question. Data analysis suggests that Spanish speakers did slightly better than other language groups on this C-Test, as shown by the means in Table No. 4.

Variable	Mean	Group
C-Test Total	114.88	All students (N=58)
C-Test Total	108.18	Arabic speakers (N=11)
C-Test Total	117.00	French speakers (N=10)
C-Test Total	120.00	Spanish speakers (N=15)
Cloze Total (E)	30.05	All students (N=58)
Cloze Total (E)	28.91	Arabic speakers (N=11)
Cloze Total (E)	30.00	French speakers (N=10)
Cloze Total (E)	30.80	Spanish speakers (N=15)
Cloze Total (A)	44.84	All students (N=58)
Cloze Total (A)	41.82	Arabic speakers (N=11)
Cloze Total (A)	46.20	French speakers (N=10)
Cloze Total (A)	44.93	Spanish speakers (N=15)

Given these data, it is uncertain whether the higher mean scores of the Spanish speakers on the C-Test total score and the Cloze total (exact scoring), and of the French speakers on the Cloze test (appropriate word scoring), are related to the native language, or whether their generally high proficiency level in English is of more significance. Because most of the other languages in the foreign student group are represented by 3 or fewer speakers, further analysis would not be productive. As Table No. 5 shows, the mean scores for the Chinese English teachers were lower on all three tests.

With respect to the question of whether the student's major field of study seems to give him/her a particular advantage in any of the test passages (either test), again, the answer appears to be negative. As in Table No. 6, the table referring to major field of study--students in particular disciplines may have some slight advantage in the Cloze test or C-Test passages representing or related to their major field of study. For example, business majors seemed to do slightly better on the marketing and economics passages in both the Cloze and the C-Test than most other students; and computer science majors seemed to do better on the computer science test passages. Further, test takers representing the other majors, including agriculture, communication, political science, English, international studies, etc., scored slightly below the mean on these particular passages.

Based on these results, I am inclined to say that the major field of study is of little significance; what is probably of more significance is the general capability that one has in reading English. In other words, I do not believe, at this point, that computer science majors, for example, had any particular advantage in those test passages related to computer science. What may be of more significance, in this instance, is the degree of familiarity that a student has with the use of computers.

I am inclined to make the same sort of intuitive generalization about the Chinese English teachers who took the C-Test and Cloze Test. Obviously, their major field of study has been English, and as indicated on the

questionnaire which each one completed, few if any had had any courses in economics, health science, computer science, business, or mathematics in the past five years. What might be of some significance, with respect to this group, is the quality of the TEFL training each has had. Unfortunately, I did not learn until after I had left China that it would have been useful to gather specific information on the types of institutions where these English teachers received their training, and on the quality and/or reputation of the programs in which they are currently teaching English as a foreign language.

I would like to turn, at this point, to an examination of the scores of the 12 test takers who scored above 50% on the Cloze test, using exact scoring. Their scores on the Cloze test ranged from 52% to 58.6% (exact scoring), and from 70.6% to 89.3% on the Cloze, using appropriate word scoring. Their C-Test scores ranged from 68.7% to 91.6%. These 12 include native speakers of Arabic, Farsi, French, Italian, Kikuyu, Spanish, and Tagalog. Their major fields of study include business administration, communications, computer science, economics, international studies, and marketing. They have all been reading in English for more than two years, and they all indicated that they frequently read English language materials for pleasure: half of them read for pleasure everyday, and the other half read English language materials for pleasure at least once a week. Most of them regard themselves as good or very good readers in their native languages, though most of them were quite modest about their reading skill in English, designating themselves as being average readers. Thus, it would seem that very good general reading skills in English mean that one can deal with, reasonably well, various types of reading content.

The last question for investigation in this small project was whether native speakers might achieve perfect or nearly perfect scores on the C-Test. As I have already indicated, the answer, in this instance, is that not one made a perfect score, and many achieved scores that I found to be surprisingly low. The mean score for native speakers was 158.17, with raw scores ranging from 135 to 176. At the same time, the average for spelling errors was 1.56; as I mentioned earlier, only 2 out of the 25 native speakers of English had no spelling errors. On individual C-Test passages, the results were as follows for native speakers (Table No. 1):

C-Test Passage	Mean Score	Range
Economics	34.20	27-37
Health	33.76	30-37
Computer Science	30.36	22-36
Marketing	27.00	18-32
Mathematics	34.20	29-37

It should be noted that the one passage on which no native speaker achieved a perfect score is the computer science passage; in general, those persons who were able to complete correctly the mutilated word chip's seemed to have had problems in completing correctly registers and/or tell; in general, the response was register and/or teach.

One might ask whether this group of educated adult native speakers is equivalent to the native speakers used by Klein-Braley and Raatz. Though they are sparing in details, I believe the two groups are equivalent. In one instance, Raatz and Klein-Braley identified the native speakers as English lecturers, English army officers, professors, and advanced university students, and they stated that the native speakers "scored at least 95% on the C-Test. (Raatz and Klein-Braley, 1981, "The C-Test--A Modification of the Cloze Procedure."). In another instance, they

asserted that a 16 year old "turns into the adult educated native speaker and achieves a perfect score on all tests." (Klein-Braley and Raatz, 1984, "A survey of Research on the C-Test.") Presumably, they mean that a 16 or 17 year old should be able to achieve a nearly perfect score on all C-Tests. Of the 25 native speakers who took this C-Test, only seven scored 95% or above--six professionals and one student. As I mentioned earlier, the scores of the remaining 18 ranged from 75% correct to 93.8% correct.

Is the C-Test--at least this C-Test--reliable? In a preliminary analysis, it would seem to be: by means of Kuder-Richardson formula 21, the reliability is estimated to be .941 for the total test, based on the mean scores for the U.S. foreign student group; for the Chinese English teachers, total C-Test reliability is estimated to be .935. Using the same formula, reliability for the individual passages ranges from .73 to .86:

C-Test Reliability KR 21

C-Test (total test)	r=.94
C-Test Econ	r=.73
C-Test Health	r=.84
C-Test Comp Sci	r=.77
C-Test Marketing	r=.84
C-Test Math	r=.86

For the U.S. foreign student group, correlations among the three sets of test scores, C-Test, the Cloze with exact word scoring, and the Cloze with appropriate word scoring are as follows:

Variable	N	Cloze E	Cloze A
C-Test Total	58	0.847	0.859
Cloze Exact	58		0.915
Cloze Approp.	58	0.915	

Correlations for the Chinese English teachers are not as strong, at least so far as the C-Test is concerned:

Variable	N	Cloze E	Cloze A
C-Test Total	28	0.569	0.664
Cloze Exact	28		0.957
Cloze Approp.	28	0.957	

Clearly, part of the difference results from the 20 point difference in C-Test mean scores for the two groups, and I would guess that that difference derives from the difference in the character of the two groups: a relatively test-wise and sophisticated group of foreign students studying in the United States, in an English speaking community, and a group of teachers for whom English is truly a foreign language, and who have never travelled outside China.

Based on this essentially exploratory investigation, I draw the following very tentative conclusions:

1. It would appear that even very difficult texts can be used for a C-Test.

2. Text neutrality may not be as necessary as Klein-Braley and Raatz have urged, though, in this investigation, the fact that the U.S. group of non-native speakers are all reasonably well educated adults, with generally good proficiency in English, may have been a contributing factor to their generally good scores.

3. C-Tests are easier to score than Cloze tests if other than the exact method is used in scoring a cloze test.

4. On the other hand, use of the exact word approach to scoring on the Cloze test is easier than scoring a C-Test, especially if the C-Test involves mutilated words of 10 or 12 or more letters.

5. C-Test scores would seem to indicate something about general reading skills, though concern with proficiency in reading does not seem to be the major concern so far as Klein-Braley and Raatz are concerned; they are very clear on this point: they view the C-Test as an alternative type of placement or general proficiency test.

This investigation has also raised a number of questions, some of which are worth further investigation. Chief among them is, what is meant by a "perfect or nearly perfect score" by educated adult native speakers? One might also ask what it is that qualifies one as an educated adult native speaker. I would also say that the matter of acceptable alternative responses, however few there might be, needs further investigation. Finally, I am not at all certain that the C-Test is a better measure than the Cloze test, especially if the exact scoring method is used. Certainly, both groups of test takers found the C-Test more frustrating, and a number of objections to its length--i.e., the number of items to be completed--were noted.

I hope some of you have found this modest investigation of sufficient interest to undertake your own C-Test investigation. I believe my handout is sufficiently detailed to enable one to begin a C-Test project. And you should know, also, that based on my correspondence with Christine Klein-Braley--which has been most helpful to me--you would find her and Ulrich Raatz more than interested in anything you might wish to undertake. Finally, because I have relied heavily on the published work of Klein-Braley and Raatz, I have tried to be scrupulously fair and accurate in presenting their ideas. Any errors or misrepresentations must be attributed to me.

BIBLIOGRAPHY

Cohen, Andrew D., Michal Segal, and Ronit Weiss Bar-Siman-Tov. "The C-Test in Hebrew." Language Testing 1, no. 2 (December 1984): 221-225*

Klein-Braley, Christine. "Advance Prediction of Difficulty with C-Tests." In Practice and Problems in Language Testing, edited by T. Culhane, C. Klein-Braley, and D.K. Stevenson. Colchester: University of Essex, 1984.

Klein-Braley, Christine. "A Cloze-up on the C-Test: A Study in the Construct Validation of Authentic Tests." Language Testing 2. no. 1 (June 1985): 76-104.*

Klein-Braley, Christine and Ulrich Raatz. "A Survey of Research on the C-Test." Language Testing 1, no. 2 (December 1984): 134-146.*

Piper, Ann. "A Comparison of the Cloze and C-Test as Placement Test Items." British Journal of Language Teaching 21, no. 1 (Spring 1983): 45-51.

Raatz, Ulrich. "Better Theory for Better Tests." Language Testing 2, no. 1 (June 1985): 60-75*

Raatz, Ulrich. "The Factorial Validity of C-Tests." In Practice and Problems in Language Testing, edited by T. Culhane, C. Klein-Braley, and D.K. Stevenson. Colchester, England: University of Essex, 1984.

Raatz, Ulrich. "Tests of Reduced Redundancy - The C-Test, A Practical Example." In Language Testing in School: AFinLA Yearbook 1985, edited by Viljo Kohonen and Antti Pitkanen. Helsinki: Pubications de L'Association Finlandaise de Linguistique Appliquee, 2985.*

Raatz, Ulrich and Christine Klein-Braley. "The C-Test; A Modification of the Cloze Procedure." In Practice and Problems in Language Testing 7, edited by T. Culhane, C. Klein-Braley, and D.K. Stevenson. Colchester, England: University of Essex, Dept. of Language & Linguistics, 1981.*

(Note: * indicates the sources consulted for this paper.)

THE KLEIN-BRALEY/RAATZ C-TEST

Christine Klein-Braley and Ulrich Raatz make the following recommendations for constructing a C-Test; where appropriate, I have added to these guidelines, based on my understanding of the procedure.

1. Use a variety of texts; the entire test should consist of 5 or 6 passages.

2. Texts for the C-Test should be relatively "neutral," so as not to give any test taker a particular advantage.

3. The entire test should have at least 100 deletions or mutilations.

4. In mutilating words, use the "rule of 2":

a) Begin with the second or third sentence and delete the second half of every other word; continue mutilations until the desired number of items has been reached.
b) The words I and a are ignored (skipped over) in the counting and mutilation process.
c) If a word consists of an uneven number of letters, the cue should contain the lesser number of letters. That is, the cue for t-h-e (3 letters) is always t_____ (one letter).
d) As I understand, names of persons should not be affected by the mutilation process.

5. The first text in the test should be relatively easy; succeeding texts should increase in difficulty.

6. Depending on text lengths, the C-Test should take between 30 and 45 minutes to administer. (In general, Klein-Braley and Raatz seem not to have set time deadlines in their testing experiments.)

7. Only exact scoring should be used for the C-Test.

8. According to Klein-Braley and Raatz, the C-Test is a norm oriented test; on average, only half the mutilations should be restored.

C-TEST TEXTS FROM KLEIN-BRALEY & RAATZ

1. English C-Test (junior version)

Once upon a time, there was a little girl who lived with her mother, who was a widow. They we___so po___that o___ day th___ had not_____ left t___ eat. T___ little gi___ went o____ into t___ woods t____ play. S____ was s___ hungry th____ she be_____ to c____. An o___ woman ca____ up t____ her. "W____ are y____ crying, m___ child?" s___ asked. "Bec____ I am s____ hungry," said the little girl. "Then you shall be hungry no more," said the old woman.

2. English C-Test (senior version)

The evening of October 30, 1938 was just like any other quiet Sunday night to most of the people of America. Many fami____ were a____ home rea____ the pap___ or conte____ listening t____ the ra____. There we____ two prog____ that ni____ which attr_____ large audi_____. One w__ the pl__ produced b____ Orson Welles. T___ listeners prep_____ themselves f___ an ho____ of comfo____ excitment, b_____, after t_____ opening announ_____, the pl____ did n____ start. Instead there was dance music.

The two texts are from Ulrich Raatz and Christine Klein-Braley, "The C-Test - A Modification of the Cloze Procedure," in Occasional Papers: Practice and Problems in Language Testing, ed. by Terry Culhane, Christine Klein-Braley, and Douglas K. Stevenson. Essex, England: Univ. of Essex, Dept. of Language and Linguistics, 1981.

3. Part of an English C-Test

Pollution is one of the big problems in the world today. Towns a____ cities a____ growing. Indu_____ is gro_____, and t___ population o____ wo_____ is gro_____. Almost every _____ causes poll_____ in so___ way o__ another. T___ air i____ filled wi___ fumes fr____ factories a_____ vehicles, a_____ there i___ noise fr__ aeroplanes a____ machines. Riv___, lakes a___ seas a____ polluted b____ factories and by sewage from our homes.

Text 3 is from Ulrich Raatz, "Tests of Reduced Redundancy - The C-Test, A Practical Example," in Language Testing in School: AF inLA Yearbook 1985, ed by Viljo Kohonen and Antti J. Pitkanen. Helsinki: Publications de L'Association Finlandaise de Linguistique Appliquee, 1985.

4. C-Test Example

There are usually five men in the crew of a fire engine. One o___ them dri____ the eng____. The lea__ sits bes___ the dri____. The ot____ firemen s___ inside t___ cab o____ the f___ engine. T___ leader h____ usually be_____ in t____ Fire Ser____ for ma____ years. H____ will kn___ how t____ fight diff_____ sorts o____ fires. S____, when t___ firemen arr___ at a fire, it is always the leader who decides how to fight a fire. He tells each fireman what to do.

Text 4 is from Christine Klein-Braley and Ulrich Raatz, "Survey of Research on the C-Test," Language Testing 1, no. 2 (December 1984): 136.

5. Archery

One Thursday afternoon the boys were doing their archery as usual. There we___ two st____ targets fi_____ yards ap____ and whe___ they ha___ shot th___ arrows a___ one, th____ had on____ to g____ to i____, collect th___, and sh____ back a____ the ot____, after fac____ about. I____ was st___ the love_____ summer wea____ and th____ had be____chicken f___ dinner, s___ that Merlyn had gone off to the edge of their shooting ground and sat down under a tree.

Text 5 is from Christine Klein-Braley, "A Cloze-up on the C-Test: A Study in the Construct Validation of Authentic Tests," Language Testing 2, no. 1 (June 1985): 83.

C-TEST (#1)

Cost-benefit analysis was developed originally with respect to water resources. Consider a multiple-purpose river development project which produces a combination of electric power, flood prevention, irrigation, and improved navigation. For po_____ output, o_____ can esti____ the pr_____ at wh____ the po_____ can b_____ sold o_____ a comme_____ basis. T_____ value o_____ flood preve_____ can b_____ judged b__ the prop_____ damage resu_____ from pa_____ floods. A_____ regards irrig_____, one c_____ compare t_____ cash va_____ of cr_____ production o_____ irrigated la_____ with produ_____ on t_____ same la_____ before irrig_____. For navig_____ benefits, o_____ can

esti_____ the prob_____ increase i_____ tonnage b_____ rail o_____ other met_____. Adding these things, one gets an estimate of total benefits from the project which can be compared with the probable cost.

Table No. 1: C-Test Scores, Native Speakers

	Mean	Range	Number
C-Test Total (out of 180)	158.174	135-176	25
C-Test Passages:			
1. Economics (37)	34.20	27-37	25
2. Health (37)	33.76	30-37	25
3. Comp. Sci. (37)	30.36	22-36	25
4. Marketing (32)	27.00	18-32	25
5. Mathematics (37)	34.20	29-37	25

Table No. 2: U.S. Foreign Students

1. Cloze Total, E (75)	30.05	15-44	58
2. Cloze Total, A (75)	44.84	25-67	58
3. C-Test Total (179)	114.88	58-164	58

Table No. 3: Passage Scores/U.S. Foreign Students

Cloze (A score):			
1. Health (14)	8.50	4-14	58
2. Marketing (13)	7.79	3-12	58
3. Economics (16)	9.65	5-14	58
4. Comp.Sci. (15)	9.94	1-15	58
5. Mathematics (17)	8.75	1-16	58
C-Test			
1. Economics (37)	26.17	10-35	58
2. Health (37)	26.46	6-36	58
3. Comp.Sci. (36)	19.81	8-32	58
4. Marketing (32)	15.74	0-29	58
5. Mathematics (37)	27.00	6-36	58

Table No. 4: Scores Grouped by Native Language

C-Test Total, all students	114.88	58
Arabic speakers	108.18	11
French speakers	117.00	10
Spanish speakers	120.00	15
Cloze Total, E, all students	30.05	58
Arabic speakers	28.91	11
French speakers	30.00	10
Spanish speakers	30.80	15
Cloze Total, A, all students	44.84	58
Arabic speakers	41.82	11
French speakers	46.20	10
Spanish speakers	44.93	15

Table No. 5: Chinese English Teachers

	Mean	Range	Number
1. Cloze Total, E (75)	28.21	15-39	28
2. Cloze Total, A (75)	42.07	26-62	28
3. C-Test Total (179)	94.11	55-156	28

Table No. 6: Scores by Major Field of Study

	Mean	Number
Cloze Marketing, A score		
All students	7.79	58
Business majors	8.11	18
Finance majors	9.67	3
Economics majors	7.67	6
All other majors	7.45	31
Cloze Economics, A score		
All students	9.66	58
Business majors	10.28	18
Finance majors	8.33	3
Economics majors	10.17	6
All other majors	9.32	31
Cloze Computer Science, A score		
All students	9.95	58
Comp. Sci. majors	11.57	7
Bus, Econ, Fin majors	9.56	27
All other majors	9.92	24
C-Test, Economics		
All students	26.17	58
Economics majors	26.50	6
Business majors	28.00	18
Finance majors	26.67	3
All other majors	26.00	31
C-Test, Computer Science		
All students	19.81	58
Comp. Sci. majors	21.43	7
Bus, Econ,Fin majors	20.33	27
All other majors	18.75	24
C-Test, Marketing		
All students	15.74	58
Business majors	17.56	18
Econ, Fin majors	16.22	9
All other majors	14.54	31

PASSAGE SOURCES -- C-Test & Cloze Test
(M.A. Hood)

C-Test Passages

1. Reynolds, Lloyd G. Microeconomics. 5th ed. Homewood, Illinois: Richard D. Irwin, Inc., 1985; page 317. (Reading level: college level, undergraduate; average sentence length = 17.71.)
2. Davis, Adelle. Let's Stay Healthy. New York: Signet, 1981; page 189. (Reading level: 9th-10th grade level; average sentence length = 14.33.)
3. "A Layman's Guide to Software." The Economist, 21 Sept. 1985, pages 86-87. (Reading level: 11th -12th grade level; average sentence length = 16.33.)
4. Hill, John S. and Richard R. Still. "Adapting Products to LDC Tastes." Harvard Business Review, March/April 1984, page 94. (Reading level: college level, undergraduate; average sentence length = 14.85.)
5. Kline, Morris. Mathematics and the Search for Knowledge. New York: Oxford University Press, 1985; page 106. (Reading level: college level, undergraduate; average sentence length = 17.33.)

Cloze Test Passages

1. Davis, Adelle. Let's Stay Healthy. New York: Signet, 1981; pages 79-80. (Reading level: 11th-12th grade level; average sentence length = 17.0.)
2. Cavusgil, S.T. and J.R. Nevin. "A Conceptualization of the Initial Involvement in International Marketing." In Theoretical Developments in Marketing. edited by C.W. Lamb, Jr. and P.M. Dunne. Chicago: American Marketing Association, 1980; page 70. (Reading level: college level, graduate; average sentence length = 14.33.)
3. Reynolds, Lloyd G. Microeconomics. 5th ed. Homewood, Illinois: Richard D. Irwin, Inc. 1985; page 334. (Reading level: 9th-10th grade level; average sentence length = 16.47.)
4. Jackson, Philip C. Introduction to Artificial Intelligence. 2nd ed., encl. New York: Dover Publications, 1985; page 281. (Reading level: college level, graduate; average sentence length = 27.57.)
5. Flegg, Graham. Numbers: Their History and Meaning. New York: Schocken Books, 1983; page 9. (Reading level: 9th-10th grade level; average sentence length = 20.3.)

NOTE: Reading level determined by means of the Dale/Chall formula as described in:
Dale, Edgar and Jeanne S. Chall. "A Formula for Predicting Readability: Instructions." Educational Research Bulletin 37 (February 18, 1948): 37-54.

TESTING LANGUAGE PROFICIENCY IN INDIA:

SOME PROBLEMATIC ISSUES

Shobha Satyanath and T.S. Satyanath

University of Pennsylvania
Philadelphia, PA

1 INTRODUCTION

This paper discusses some important issues associated with language testing in multilingual and pluricultural India. It attempts to cover some of the pedagogic and sociolinguistic issues confronting language testing in India. This is not to suggest that all problematic issues have been taken care of here. In addition, some problems are general in nature and refer to different languages and some are specific only to Hindi. Although the discussion centers around Hindi, it applies equally well to other Indian languages. We have tried to refer to other language situations wherever it is relevant.

The situation of language testing in India is quite complex because of several reasons. On the one hand there is a paucity of developed language tests and a blind adaptation of existing tests from the western countries raises certain pedagogical issues. At the same time the peculiar sociolinguistic situation of India on the other hand raises certain sociolinguistic issues. We have attempted to point out these issues and have suggested that the development of language testing instruments has to be taken up in the light of the issues discussed here.

2 PEDAGOGIC ISSUES

2.1 Any language should take into account the teaching methodologies used in the teaching program and the conditions under which a language is being learnt. As far as language teaching is concerned, certain considerations are of great importance. These considerations are as follows: (i) whether the language concerned is taught as a language or as a subject; (ii) whether the emphasis is on mastery of formal aspects or on communicative skills; and (iii) how much time is being spent on language teaching.

2.1.1 As far as language learning conditions in India are concerned, a language can be learnt broadly under two conditions, namely, acculturation or enculturation situations. This is true whether it is being learnt as a first or as a second language. In acculturation situations, the target language is also the language of the majority community. This reinforces learning through a rich language environment. The enculturation situation, on the other hand, is one where the target language is not the language of the majority community. Hence, it lacks the environmental reinforcement

TABLE 1: DETAILS OF LEARNING OF HINDI AS FIRST AND SECOND LANGUAGE IN ACCULTURATION AND ENCULTURATION SITUATIONS

	First Language	Second Language
Acculturation	Hindi in Hindi speaking states for Hindi MT speakers	Hindi in Hindi speaking states for non-Hindi MT speakers
Enculturation	Hindi in non-Hindi speaking states for Hindi MT speakers	Hindi in non-Hindi speaking states for non-Hindi MT speakers

and is learnt in a limited environment either at home or in school through formal teaching. Table 1 gives the details of Hindi being learnt as a first and second language in acculturation and enculturation situations in India. However, it needs to be pointed out that learning a language in enculturation situations creates more problems to the learner than the acculturation.

2.1.2 In India, Hindi is being taught in schools all over the country following the 'three language formula'. As a result, depending upon which state one considers, Hindi is studied compulsorily by every student as first, or second, or third language for a period of ten, or six, or three years respectively (cf. Table 2). Besides, Hindi is the official language of the nation, one of the link languages among different states and also, the official language of seven states (as recognized in the Constitution of India). So far as learning of Hindi is considered one finds both enculturation and acculturation situations in the country. The enculturation situation exists in all the non-Hindi speaking states where the target language group is not present and the language is learnt only in the formal class room context for a few years. The acculturation on the other hand, is found in the regions where Hindi and its dialects are spoken. The example of the latter is found in Delhi, Madhya Pradesh and Bihar, for instance. Here, the target language is in fact the language of the majority community. In acculturation situations, the environment and the mass media reinforce the learning of Hindi, whereas in an enculturation situation, they are conspicuously absent. However, the television media and the Hindi films do reinforce the learning to some extent, although this type of reinforcement is confined only to receptive skills and not to productive skills.

As far as teaching is concerned, languages are taught in schools like any other subject - mathematics, geography or science and not so much as a language. The emphasis is on reading and writing abilities, knowledge of literature from ancient to modern times and on grammar. This is in contrast with intensive language teaching programs in the education of the western countries. As a matter of fact, in many parts of rural India, teaching a language through the rote system still continues, where even the children in standards II and III cannot read properly, but can reproduce the entire lessons from their text books by heart without mistakes.

2.1.3 The difference in language teaching programs between India and the west is basically due to the educational policies as well as the nature of multilingualism that exists in these countries. A glance at the language teaching programs reveal that the divergence and variance in the existing

TABLE 2: POSITION OF HINDI IN CURRICULUM IN SOME SELECTED STATES IN INDIA

Sl. No.	State	Major medium	Subordinate medium	Principal official language	Language used in official	POSITION OF HINDI IN SCHOOLING				
						Major first lang.	Major second lang.	Minor second lang.	Compulsory third language	Official language
1	Andhra Pradesh	Te	E.H.U.Ma,K	Te/U	E,Te,U,H	I-X	V-X	-	-	-
2	Assam	A	E, B.H,U	A/B	E,A,B,H	I-XI	-	-	IV-XI	-
3	Kerala	M	E,Ta,K	M	E,M	-	-	-	V-X	-
4	Karnataka	K	E,H,U,Ma, Ta,Te	K	E,K	-	-	V-X	VI-X	-
5	Orissa	O	E,H	O	E,O	I-XI	-	-	VI-XI	IX-XI
6	Tamilnadu	Ta	Te,M,K,U, H,E,G	Ta	E,Ta	I-XI	-	-	-	-
7	West	B	E,H,U,Ne,O, Te,G,A,L,M, Ma,Mt,P,Sa, Sd,Ta	B	E,B,Ne	I-XI	-	-	VI-Xi	-
7	Delhi	H	E,U	H	E,H	I-XI	V-XI	-	-	-

A: Assamese, B: Bengali, E: English, G: Gujarati, H: Hindi, K: Kannada, M: Malayalam, Ma: Marathi, Mt: Maithili, Ne: Nepali, O: Oriya, P: Panjabi, Sa: Santali, Ta: Tamil, Te: Telugu, U: Urdu

(Computed from Tables I, III, IVA, V, VI, and VIII of Chaturvedi and Mohele 1976)

testing instruments is directly in correlation with the divergence in the teaching programs, which cater to different needs of these societies. For instance, the assessment approach, which assesses the communicative abilities in a variety of contexts and situations is a direct outcome of the bilingual education and/or bilingual immersion programs. These assessment approaches assess child's linguistic behaviour not only in the classroom, but also in the playground, at home and in peer-group settings. The major aim is to assess the developing skills of the children in semifield conditions who come from divergent and different linguistic backgrounds. On the other hand, the standardized language tests measure academic skills at different levels in intensive language teaching programs. Similarly the Council of Europe Crediet Scheme which measured the linguistic proficiency on a three-point scale, from the tourist level to the threshold level, also reflected the nature of its language teaching program. The idea behind the self-assessment scale is to make the learner aware of his skills at each level, so that he could devise his own strategies to learn better.

2.1.4 These considerations of teaching methods and learning conditions are important from the point of testing because testing is an integral part of learning and teaching. Different types of teaching and learning situations may result in different kinds of learning, both in terms of degree and variety. Also, equally important are the purpose of testing and the use that one is going to make once the data is obtained from the tests. For example, one has to find out whether it is appropriate to test one's communicative skills in a language when the emphasis in the teaching program is only on teaching the formal aspect of the language concerned. Similarly, is it appropriate to test the formal academic skills, where the result is actually the colloquial variety and just the interpersonal skills? In fact, teaching goals may drastically differ from program to program not only from one country to another but also within a country. It is in this respect that one finds the Indian society drastically different from the west.

2.2 This section highlights some general problems that arise while adapting the existing tests for Indian situations. These include (i) time factor involved in testing, (ii) unfamiliarity and alien nature of the tests and (iii) economic aspects of the tests.

2.2.1 If the time taken to administer the test is very long, there is a chance that the attention span of the students may be reduced. This will surely affect the reliability of the test. If the test takes a long time to administer, then to standardize such a test would be extremely difficult. In addition, tests of long duration are strenuous and also torturous to the test taking students. The Indian rural population has been brought up in a tradition where the concept of testing is not really deep rooted. It is almost impossible to administer tests of long duration on the rural Indian population.

2.2.2 Unfamiliarity with the gadgets of the test may cause a shock to the learner and may prevent the student from presenting a true performance. Labov (1969) has clearly pointed out the observers' paradox that creeps into the study due to the asymmetric relationship between interviewers and informants in an interview situation. With regard to testing, the formality associated with the testing procedure and the unfamiliarity with the test gadgets together may create problems.

One of our experiences in rural parts of Madhya Pradesh illustrates how the time factor and the unfamiliarity of the test procedure created

the problems in testing*. A cloze test in Hindi was administered to the secondary school students in Hoshangabad district. The students of the central school (a chain of schools all over India with the modern technical knowhow available to them), who were familiar with the test procedures finished the test with a mean time of 7.78 minutes whereas, the students of state secondary board school took as much as 13.51 minutes. When the same test was administered in a tribal area, students took much longer time (boys took a mean time of 24.50 minutes and girls 33.71 minutes). Some of them even gave up their hope to complete and returned the test. The general proficiency level turned out to be very low. More than 72% of the children are at the frustration level. This is particularly interesting because Hindi is the medium of instruction in all these schools.

2.2.3 Finally, the economic aspects of a test plays a significant role in choosing a test not only in India but also in other third world countries. On the one hand there are those huge and expensive test batteries, and an abnormally large population on which they have to be administered. On the other hand, the economic implication of their use may discourage even the Government agencies to undertake the work of preparation of such test batteries. The above discussion indicates the inadequacy of the existing tests in the west for the Indian situation and the need for tests which are smaller in size, less expensive, take less time to administer, are simple to administer and familiar to the students.

3 SOCIOLINGUISTIC ISSUES

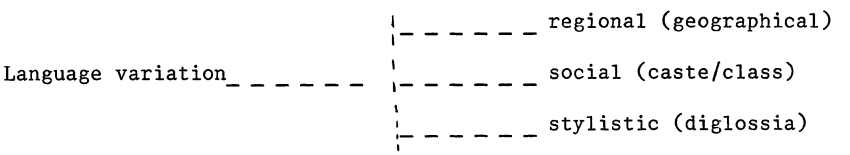
3.1 Among sociolinguistic issues that are discussed here, some are general ones and apply to all the Indian languages and some are specific only to Hindi. However, all of them have significant relevance to language testing. In the following paragraphs, we will discuss language testing in Indian context with reference to multilingualism and pluriculturalism and language variation.

3.2 The very nature of multilingualism in India officially being backed by the language planning policy of the Government might be advantageous (cf. Gumperz 1958, Srivastava, 1976) or disadvantageous or a social burden (cf. Chickermane, 1971, Ross, 1965). The latter viewpoint has been more or less refuted now. But the fact that more than one language exists in the curriculum raises the question whether proficiency in one language should be tested with reference to other languages in the curriculum. In other words, whether there is a correlation between proficiency in one language and proficiency in other languages in a multilingual education programme? In a multilingual setup, different languages perform different functions. Besides, attitudes and motivation of a particular speaker towards the languages he uses and the attitudes of the speech community towards the speaker may have significant bearing on linguistic competence. It needs to be pointed out that in India not only the patterns of language use but also attitude towards languages and speech communities along with other learner variables have to be incorporated into language testing. Satyanath and Agnihotri (1983) have suggested the need to outline the sociolinguistic profile of the speech community being studied prior to the design of educational programs.

3.3 Language variation: Although language variation is a characteristic of languages all over the world, South Asia has been considered as an area with a high degree of variation in the languages being used there. Gumperz

*All data regarding rural parts of Madhya Pradesh are from a project on Evolving Hindi Teaching Materials to Rural Children in Madhya Pradesh, a project with which we were associated. The project was sponsored by Eklavya, E1/208, Arera Colony, Bhopal 462016, India.

(1958) has shown that Kalapur, although a small village with a small area of interaction, demonstrates a complex structure of linguistic variation which could only be explained in terms of a sociolinguistic background. Broadly speaking, the language variation in the Indian context can be studied under three main heads shown below:



3.3.1 In the course of the history of the Hindi language we have seen that one of its regional varieties developed as the standard variety and that different regional varieties were considered the standard variety at different points of time. This multiplicity of standards in the tradition of the Hindi language has actually encouraged dialects like Maithili and Bhojpuri (hitherto accepted as dialects of Hindi by Hindi protagonists) to stand up and threaten to establish their own identity as separate languages. The fact that some of these dialects have a rich culture and literary tradition and are linguistically distinct (cf. Singh 1986 for such differences) so as to provide problems in mutual intelligibility further complicates the Hindi situation. This has led to a parochialization tendency which has reinforced the regional features of Hindi, particularly phonological, morphological and lexical to be nurtured well so as to strengthen and maintain the differences between the dialects and the standard Khariboli. This, in addition to the delay in the standardization process of Khariboli has posed problems not only in teaching but also in testing. The problem may not be acute as far as the receptive skills are concerned but coming to productive skills, the intra-lingual variation has given rise to a number of phonological and morphological deviations. A look at the data in Table 2 would demonstrate the differences between standard Hindi and the variety spoken in Hoshangabad. How to take into consideration these deviations and what sort of weightage has to be given to them are some of the serious questions confronting language testing.

3.3.1.1 There are a number of deviant varieties of Hindi in the non-Hindi speaking states. Although the Hindi spoken by the other Indo-Aryan language speakers does not have several phonological variables, the south Indian variety of Hindi is marked by its phonological, morphological and syntactic deviations. Narasimharao and Jaswal (1976) and Koshal (1976) have shown some of these deviations in the writings of creative writers and translators from south India with Kannada, Malayalam, Tamil and Telugu as their mother-tongues. The case of Hindi text-books in the Hindi speaking areas is also not very satisfactory. Lack of consistency in the case of orthographic symbols used in the text and the use of regional variants have together added to the complexity of the situation. In addition the texts are not prepared scientifically. A study is in progress where we have pointed out these issues through an analysis of class I text-book prescribed by the Board of Secondary Education of Madhya Pradesh. Testing the language of the learners under such circumstances have to take into account the variation built in within the text-books.

3.3.2 The social variation in South Asia has been observed to be due to the very nature of its social organization, namely, caste. Most of the studies done on social variation of language in South Asia are studies in caste variation. In fact, there are scholars (Pandit 1969, Pattanayak 1975) who have voiced serious objections to the over-domination of the study of caste dialects. It appears that language variation along caste lines is more marked in Dravidian than in Indo-Aryan languages. It has already been pointed out that there exist clearly remarkable phonological,

morphological and lexical differences between these caste dialects (cf. McCormack 1960, Bright 1970, Bright and Ramanujan 1962). Although such differences do not exist in the case of Hindi, this fact raises some general questions: what are the implications of testing the standard variety which is always closer to the most prestigeous variety - either the Brahmin or any other dominant caste's speech style? How to take care of a situation when in a society a particular group has an advantage and the rest are in a disadvantageous position? In fact, the latter question is being asked by the exponents of non-Brahmin movements in South India and Maharastra. It is a fact that the South Asian society is capable of living with all these inbuilt variations. But at the same time, it should be pointed out that no systematic study, so far, has been made to point out the role of caste dialects in educational achievement and language testing in particular.

3.3.2.1 Apart from caste, the urban-rural parameter appears to have some disparity with language testing. It appears that rural children are in a disadvantageous position compared to the urban ones. Although there does not seem to be any empirical evidence for this claim it needs to be pointed out that during our field work in Madhya Pradesh, most of the rural children who were administered cloze tests to measure general proficiency were found to be at the frustration level (about 72%), in spite of the fact that the tests were designed from a large variety of texts similar to the kind prescribed for these children. In addition, they also lacked attention span and took an abnormally long time to complete it (refer to 2.2.2). All these things point out the fact that the existing mode of testing might have placed the urban children at an advantageous position as compared to their rural counterparts. Whether there is a need to train the rural children to acquaint themselves with the existing techniques or whether one should evolve new techniques to suit the rural children is a matter to be considered with great care.

3.3.2.2 Like the urban-rural divide, sex has also shown to be a parameter which puts the boys in an advantageous position. Although most of the sociolinguistic studies done abroad as well as in Indian urban centres (Trudgill 1974, Agnihotri 1979, Mukherjee 1980, Satyanath 1982) have shown that women assimilate language faster and maximally than men, the case of rural girls appears to be violating this norm. Satyanath et al. (1984) found that with the reading comprehension tests they administered for the rural children of Madhya Pradesh, the boys did significantly better than the girls. They observed that the poorest performance was by the girls of higher age-group. This they ascribe to the social restrictions on girls in the traditional Indian setup, late entry to school and to other reasons. It is not only that the girls face increasingly more restrictions as they grow older, they also have to share considerable amount of work load at home. In fact, the traditional Indian society puts a very high premium on the education of boys and neglects that of girls. How to take care of this anomaly in testing is a question which is difficult to answer at this stage.

3.3.3 The stylistic variation which is exemplified by diglossia, appears to be more a characteristic feature of Dravidian languages and Bengali. The two styles, literary and colloquial or written and spoken have their own distinct phonological and morphological rules. In fact, the differences between the literary and colloquial styles are so large that internalization of literary variety even takes several years of schooling (McCormack 1960). However, in the case of Bengali, Singh and Maniruzzuman (1983) and Singh (1986) point out that the structural differences between the two codes is not great.

Although De Silva (1974) shows that literary variety of Sinhala is easily comprehensible indicating that the entire speech community might

have the receptive skills but not the productive ones, Shanmugam Pillai (1960) points out that attaining literacy in a highly diglossic language presents more problems than languages without such stylistic clevage. Since Hindi is a non-diglossic language, one may think this does not concern Hindi teaching and testing. But one has to look at the fact that the Bengali diglossic situation (Dimok 1960) was actually a development of eighteenth century Bengal with the growth of Calcutta and the role of Fort William College in the standardization of Bengali. While the two styles in Bengali have a strong correlation with each other and the distinction is on the verge of slowly disappearing in the case of Bengali in West Bengal (Singh 1976), but still growing strong in Bangladesh (Singh and Maniruzzuman 1982), it has significant implications for Hindi. In fact, Hindi which historically lacks diglossia, seems to be developing a diglossic situation through inventing a Sanskritised variety to fill in the gap. Another question that arises here is which variety to be tested--the written or spoken.

4. To summarize, we have raised the following issues:

(i) Whether the existing language tests are appropriate for the Indian situation?
(ii) Can we modify the existing tests to suit Indian population or do we need to develop new tests?
(iii) Do we need different tests for enculturation and acculturation situations?
(iv) Should a test always be norm referenced or criterion referenced?
(v) If different learning situations reinforce different varieties, is it appropriate to test always the standard variety only?
This cautions against indiscriminate use of existing language tests which have been designed for and standardized for western population. At the same time, it urges for a need to look for fresh solutions, based on the nature and requirements of specific countries.

REFERENCES

Agnihotri, R.K. 1979. Processes of Assimilation: A Sociolinguistic Study of Sikh Children in Leeds. Unpublished D. Phil. dissertation, University of York.

Bright, W. 1970. Phonological Rules in Literary and Colloquial Kannada. Journal of American Oriental Society 90.1: 140-44.

_____ and A.K. Ramanujan. 1962. Sociolinguistic Variation and Language Change. Proceedings of the Ninth International Congress of Linguists. ed. by H.G. Lunt: 1107-13. The Hague: Mouton.

Chaturvedi, M.G. and B. V. Mohele. 1976. The Position of Languages in School Curriculum in India. New Delhi: NCERT.

Chickeremane, D.V. 1971. Impact of Bilingualism on the Progress of Children in Primary Schools in Rural Areas. New Delhi: NCERT.

DeSilva, M.W.S. 1974. Convergence in Diglossia: the Sinhalese case. Contact and Convergence in South Asian Languages. ed. by F.C. Southworth and M.L. Apte. IJDL 3.1: 60-91.

Dimok, E.C. 1960. Literary and Colloquial Bengali in Modern Bengali Prose. Linguistic Diversity in South Asia. ed. by C.A. Ferguson and J. J. Gumperz, International Journal of American Linguistics 26.3, Part 3: 43-63.

Gumperz, J.J. 1958. Dialect Differences and Social Stratification in a North Indian Village. American Anthropologist 60.4: 668-82.

Koshal, S. 1976. Kannada-Hindi: Characteristics of Hindi written by Kannada Speakers. Studies in Bilingualism. ed. by B.G. Misra. Mysore: CIIL.

Labov, W. 1969. The Logic of Non-standard English. Language and Social Context. ed. by P.P. Giglioli. 1972. Harmondsworth: Penguin.

Mukherjee, A. 1980. Language Maintenance and Language Shift among Punjabis and Bengalis in Delhi: A Sociolinguistic Perspective. Unpublished Ph.D. thesis, University of Delhi.
Narasimharao, K.V.V.L. and S.C. Jaswal. 1976. Kannada-Hindi Studies in Bilingualism. ed. by B.G. Misra. Mysore: CIIL.
McCormack, W. 1960. Social Dialects in Dharwar Kannada. Linguistic diversity in South Asia. ed. by C.A. Ferguson and J.J. Gumperz, International Journal of American Linguistics 26.3, Part 3.
Pandit, P.B. 1969. Parameters of Speech Variation in an Indian Community. Language and Society in India: Transactions. Simla: Institute of Advanced Studies.
Pattanayak, D.P. 1975. Caste and Language. IJDL 4.1: 97-104.
Ross, A.D. 1965. Some Implications of Multilingualism. Towards a Sociology of Language. ed. by J.K.N. Unnithan, Indra Deva and Yogendra Singh. New Delhi: Prentice Hall.
Satyanath, Shobha. 1986. Evolving Reliable Tests for Testing Proficiency in Hindi as a Second Language. A comprehensive report submitted to the Central Institute of Indian Languages, Mysore. University of Delhi.
_____, T. S. Satyanath and R. K. Agnihotri. 1984. On Reading Comprehension among Rural Children. Paper read at the XIII All India Conference of Linguistic Society of India held at CIIL, Mysore, May 29-31, 1984.
Satyanath, T.S. 1982. Kannadigas in Delhi: A Sociolinguistic Study. Unpublished M. Phil. dissertation. University of Delhi.
_____and R. K. Agnihotri. 1983. Teaching Native Language in Non-native contexts: The case of Kannada in Delhi. Paper read at Seminar on Teaching/Learning of Kannada. CIIL, Mysore.
Shanmugam Pillai, M. 1960. Tamil: Literary and Colloquial Linguistic Diversity in South Asia. ed. by C.A. Ferguson and J. J. Gumperz. International Journal of American Linguistics 26.3, Part 3: 27-42.
Singh, U.N. 1976. Diglossia in Bengali: A Study of Attitudes. Papers in Linguistic Analysis 1.1: 94-99.
_____. 1986. Some aspects of Maithili-Hindi Divide: Evidence from Syntax. Paper read at the Seminar on Contrastive Linguistics (with special reference to Hindi) held on March 13-15, 1986 at the Central Hindi Institute, Delhi.
_____and Manirazzuman. 1982. Diglossia in Bangladesh and Language Planning. Calcutta: Gyan Bharati.
Srivastava, R.N. 1976. Bilingualism: Myth or Reality. Indian Bilingualism ed. by G. Sharma and Suresh Kumar. Agra: Kendriya Hindi Sansthan.
Trudgill, P. 1974. The Social Stratification of English in Norwich. Cambridge: Cambridge University Press.

INDEX