

**GEORGETOWN UNIVERSITY ROUND TABLE
ON LANGUAGES AND LINGUISTICS**

*Language Teaching, Testing, and Technology:
Lessons from the Past with a View Toward the Future*

1989

James E. Alatis, *Editor*

GEORGETOWN UNIVERSITY ROUND TABLE ON LANGUAGES AND LINGUISTICS 1989

*Language Teaching, Testing, and Technology:
Lessons from the Past with a View Toward the Future*



James E. Alatis, *Editor*

Georgetown University Press, Washington, D.C.

Bibliographic notice

Since this series has been variously and confusingly cited as: *Georgetown University Monograph Series on Languages and Linguistics*, *Monograph Series on Languages and Linguistics*, *Reports of the Annual Round Table Meetings on Linguistics and Language Study*, etc., beginning with the 1973 volume the title of the series was changed.

The new title of the series includes the year of a Round Table and omits both the monograph number and the meeting number, thus: *Georgetown University Round Table on Languages and Linguistics 1988*, with the regular abbreviation *GURT '89*. Full bibliographic references should show the form:

Lado, Robert. 1989. Acquisition vs. learning in reading pronunciation by adult EFL students. In: *Georgetown University Round Table on Languages and Linguistics 1989*. Washington, D.C.: Georgetown University Press. 000-000.

Copyright ©1989 by Georgetown University Press
All rights reserved
Printed in the United States of America

Library of Congress Catalog Number: 58-31607
ISBN 0-87840-124-5
ISSN 0196-7207



To WILGA M. RIVERS
TEACHER, AUTHOR, SCHOLAR
IN RECOGNITION OF HER SELFLESS DEDICATION
TO THE TEACHING OF LANGUAGES

Contents

The Georgetown University Bicentennial and the fortieth Round Table on Languages and Linguistics

James E. Alatis, <i>Dean, Georgetown University School of Languages and Linguistics and Chair, Georgetown University Round Table on Languages and Linguistics 1989</i>	
Welcoming remarks	1
Charles A. Ferguson, <i>Stanford University</i>	
On behalf of the Center for Applied Linguistics	5
Charles L. Currie, S.J., <i>Director, Bicentennial Celebration</i>	
On behalf of the Georgetown University Bicentennial	7
James E. Alatis, <i>Chair, Georgetown University Round Table on Languages and Linguistics 1989</i>	
Presentation of Bicentennial Medals to Henry and Renée Kahane	9
Sir John Lyons, LittD., F.B.A., <i>Cambridge University</i>	
The last forty years: Real progress or not?	13
<i>Language teaching</i>	
G. Richard Tucker, and JoAnn Crandall, <i>Center for Applied Linguistics</i>	
The integration of language and content instruction for language minority and language majority students	39
Veronica Lawson, <i>London, England</i>	
Practice makes less imperfect: Users' needs and their influence on machine translation development	51
Robert Lado, <i>Dean emeritus, Georgetown University</i>	
Acquisition vs. learning in reading pronunciation by adult EFL students	59
Claire J. Kramsch, <i>Massachusetts Institute of Technology</i>	
Discourse and text: A narrative view of the foreign language lesson	69
Charles A. Ferguson, <i>Stanford University</i>	
Language teaching and theories of language	81

Muriel Vasconcellos, <i>Pan American Health Organization</i> Cohesion and coherence in the presentation of machine translation products	89
Richard Lutz, <i>Georgetown University</i> Second language acquisition: Do we really want a unified theory?	106
Karin C. Ryding, <i>Georgetown University</i> Less commonly taught languages: The current situation	114
Peter Strevens, <i>Wolfson College, Cambridge and the University of Illinois</i> Towards a rationale for language teaching technology	122
Ali Hajjaj and Balkees Al-Najjar, <i>Kuwait University</i> ESL program evaluation: Realities and perspectives	133
Frederick Bosco and Anna De Meo, <i>Georgetown University</i> Discourse frames and the cycle of instruction	142
Ross Steele, <i>University of Sydney</i> Teaching language and culture: Old problems, new approaches	153
Henry Kahane, <i>University of Illinois at Urbana-Champaign</i> The role of language in the immigrant's life	163
Maria Ibba, <i>Università Cattolica del Sacro Cuore</i> New trends in foreign language teaching: Teaching English in the Italian medical school curriculum	167
Irene Thompson, <i>George Washington University</i> Lexical search strategies in L2: A developmental analysis	177
Wilga M. Rivers, <i>Harvard University</i> Interaction and communication in the language class in an age of technology	186
 <i>Language testing</i>	
Ray T. Clifford, <i>Defense Language Institute</i> Technological, methodological, and assessment challenges: Can the foreign language teacher survive?	198
John L. D. Clark, <i>Defense Language Institute</i> Multipurpose language tests: Is a conceptual and operational synthesis possible?	206

Peter H. Lowenberg, <i>Georgetown University</i>	
Testing English as a world language: Issues in assessing nonnative proficiency	216
Richard T. Thompson, <i>Center for Applied Linguistics</i>	
Oral proficiency in the less commonly taught languages: What do we know about it?	228
J. Michael O'Malley, <i>Evaluation Assistance Center, Georgetown University</i>	
Language proficiency testing with limited English-proficient students	235
Thea C. Bruhn, <i>Foreign Service Institute</i>	
'Passages': Life, the universe, and language proficiency assessment	245
Rebecca M. Valette, <i>Boston College</i>	
Language testing in the secondary schools: Past experience and new directions	255
Heidi Byrnes, <i>Georgetown University</i>	
Who is in charge in the learner-curriculum-testing connection?	265
 <i>Language technology</i>	
Sue K. Otto, <i>University of Iowa</i>	
Assessment, articulation, accountability: New roles for the language lab	276
Nina Garrett, <i>University of Illinois at Urbana-Champaign</i>	
The synergism of technology and theory in classroom second language acquisition research	288
Michael Zarechnak, <i>Georgetown University</i>	
Semantic subclasses of temporal nouns	295
Protase E. Woodford, <i>Educational Testing Service</i>	
From wire recorder to satellite dish: The impact of technology on language teaching	304
June K. Phillips, <i>Tennessee Foreign Language Institute</i>	
Challenging teachers and harnessing technology	312
Frank Otto, <i>Brigham Young University</i>	
Building on the past: New directions in CAI/IL	320
Judith G. Frommer, <i>Harvard University</i>	
Language learning, cultural understanding, and the computer	332

Sergei Nirenberg, <i>Carnegie-Mellon University</i> New developments in knowledge-based machine translation	344
Ricardo M. Paiva, <i>Georgetown University</i> Culture in the language class: Videos to bridge the gap	358
Margaret King, <i>Université de Génève</i> Using unification grammars for analysis and synthesis	364
Reinhold Freudenstein, <i>Philipps - Universität, Marburg</i> Testing and technology in Germany revisited: What is left? What can be hoped for?	372
Makato Nagao, <i>Kyoto University</i> New directions of machine translation	378
Winfred P. Lehmann, <i>University of Texas at Austin</i> Machine translation: Achievements, problems, promise	385
Stephen Krashen, <i>University of Southern California</i> Language teaching technology: A low-tech view	393
G. Richard Tucker, <i>Center for Applied Linguistics</i> Michael Zarechnak, <i>Georgetown University</i> Appendix: Three historical notes	408

Welcoming remarks

James E. Alatis

Dean, Georgetown University School of Languages and Linguistics

*Chair, Georgetown University Round Table on Languages and Linguistics
1989*

Good evening, ladies and gentlemen. My tasks this evening are simple and happy: to welcome, to thank, to announce, and to introduce.

1. Welcome to Georgetown University, to our School of Languages and Linguistics, and to the Georgetown University Round Table on Languages and Linguistics 1989. We are delighted that since 1950 we have been able to host scholars and teachers interested and committed to languages and linguistics. Our colleagues in this year's *Round Table 1989* will discuss *Language Teaching, Testing and Technology: Lessons from the Past with a View to the Future*.

A word of welcome also to those who participated in the Round Table pre-sessions. These pre-session meetings allow us all to include a wide range of quite specialized topics which a more general conference could hardly hope to address. We are delighted to provide a setting for such groups, some of which, such as the Interagency Roundtable, have met here for many years. We are grateful to the organizers of these pre-sessions and congratulate them.

2. Now let me thank all of you who have come: speakers, scholars, teachers, students. This is our fortieth annual Round Table meeting. We are grateful that you could join us for this anniversary and we welcome your participation.

I would like to express my special gratitude to Mrs. Carol Kreidler, the conference coordinator, who helped identify and invite to our meeting so many of the finest scholars concerned with various aspects of language study. Unfortunately she is not here this evening; for as a member of the Executive Board of TESOL, she was committed to attend their annual conference in San Antonio, and I insisted that she honor that commitment.

But her assistant coordinator, Mrs. Jackie M. Tanner, Director of Language Learning Technology in the School of Languages and Linguistics, is here together with Round Table assistants Susan Huss and Laura Klos Sokol, Fellows in the Linguistics Department. Let me express to them our joint gratitude for the assistance I know they will provide during the next few days.

3. My first announcement is that this year's Round Table meeting again coincides with the National Foreign Language Week proclaimed by President George Bush. I am sure you share with me gratitude for his gracious words

2 / Georgetown University Round Table on Languages and Linguistics 1989

and thoughts about the importance of foreign language learning and its contribution to our cultural enrichment.

My next announcements are about anniversaries. This year, 1989, marks the two hundredth anniversary of the ratification of the Constitution of the United States of America, the two hundredth anniversary of the inauguration of our first American President, George Washington, and the two hundredth anniversary of the founding of Georgetown by the first American Catholic bishop, Archbishop John Carroll. These are bicentennials we announce and celebrate with true joy.

This year is also the fortieth anniversary of the founding of the School of Languages and Linguistics and this is our fortieth Georgetown University Round Table on Languages and Linguistics.

During this bicentennial year, I have been privileged to award a Bicentennial Medal to the Mexican poet, Octavio Paz, when he kindly gave a reading of his poetry here this year. And in recognition of our long-term relationship with the French Embassy, I was able to bestow a similar medal on the French Ambassador, His Excellency M. de Margerie.

This year also marks the thirty-fifth anniversary of the Georgetown University-IBM Russian/English English/Russian Machine Translation Project. Professor Michael Zarechnak, a member of the original project, has arranged some papers to commemorate this event.

This year, moreover, is the thirtieth anniversary of the Center for Applied Linguistics (CAL), originally an independent agency of the Modern Language Association and originally funded by the Ford Foundation. All of us commend the Center for its dedicated service to applied linguistics. During those thirty years Georgetown University and the Center for Applied Linguistics have enjoyed and benefited from a mutual, amicable, and working relationship. We look forward to many years of continued friendship and productive work with the Center.

My final announcement is that this year is not only the centennial of Australia but includes the seventieth birthday of one of our favorite Australians, Professor Wilga Rivers of Harvard University. In recognition of her many contributions to the study of language and language teaching we are dedicating to her this volume, the Georgetown University Round Table on Languages and Linguistics 1989.

4. My fourth task consists of three introductions.

First, in view of our extensive commitment to applied linguistics, we felt it appropriate to extend an invitation to Professor Charles A. Ferguson, the first Director of the Center for Applied Linguistics, to address this conference. A distinguished scholar, now a professor at Stanford University, we are honored to have him as our guest tonight.

Second, it is my pleasure to introduce to you the Reverend Charles L. Currie, S.J., Director of the Georgetown University Bicentennial Celebration.

Third, on this fortieth anniversary of the Georgetown University Round Table on Languages and Linguistics, I have invited the eminent British linguist, Sir John Lyons, to give the opening address. I feel deeply honored to have this opportunity to thank him for having agreed to come.

You who have read his works recognize the enormous influence he has had on the study of language and linguistics. His publishing career spans a quarter century, beginning in 1963 with the publication of *Structural Semantics*. Other works such as his *Introduction to Theoretical Linguistics*, *Noam Chomsky and New Horizons in Linguistics* have been translated into numerous languages, including Dutch, French, German, Italian, Japanese, Polish, Portuguese, Russian, Spanish, and Swedish. He is a frequent contributor to such publications as the *Journal of Linguistics*, numerous scholarly journals, and *The Times Literary Supplement*. To describe the influence of John Lyons as truly worldwide is not hyperbole, but fact. His two-volume work, *Semantics*, is recognized as the most definitive treatment of the subject to date.

Born in 1932, John Lyons was educated at St. Bede's College in Manchester and Christ College at Cambridge. His distinguished teaching career has included posts at the Universities of London, Cambridge, Edinburgh, and Sussex. He is currently Visiting Professor of Linguistics at the University of Sussex and Master of Trinity Hall at Cambridge. He has received honorary degrees from the Université catholique de Louvain, the University of Reading, and just last year, the University of Edinburgh. He is a fellow of the British Academy and an honorary member of the Linguistics Society of America. In June 1987, he was knighted for services to the study of linguistics.

The topic of Sir John's address this evening is 'The Last Forty Years: Real Progress or Not?' It is a great honor for me to welcome him to the podium.

On behalf of the Center for Applied Linguistics

Charles A. Ferguson
Stanford University

As the first Director of the Center for Applied Linguistics all of you certainly have my sincere congratulations and warm wishes for creative growth during the third century opening up for your University. At first I thought I was being asked to speak because your Dean and I are old friends, friends for more than forty years--I even forgot the exact date of our first acquaintance, so long have I known him. But on reflection I decided that his choice was just, for I have had the unique good fortune of being present here in Washington, D.C. at the beginnings of all sorts of events which have had to do with the language and linguistic community.

I was here in 1946 at the Foreign Service Institute of the Department of State when the section that deals with language teaching and language research was beginning. I was in Washington when Father Edmund A. Walsh and Leon Dostert began the then Institute of Languages and Linguistics, now School of Languages and Linguistics. I remember those of us that were here at the time thought, 'That's a fine idea'; but we had no idea that it would finally develop the way that it did. Forty years ago we never could have predicted that it would grow, diversify, and become a leading, significant institution on the American scene.

I was even here at the first Round Table. We met in the old Multilingual Room, which, unusual at that time, had provisions for simultaneous interpreting. Incidentally, the table was not 'round', as the folklore has it, but 'U'-shaped, with sharp edges, but we could all easily fit around it. But that first Georgetown Round Table on Languages and Linguistics which I attended was a little unlike the Institute which sponsored it, or, for that matter, the Foreign Service Institute. In both of them one could foresee a kind of potential bureaucratic persistence for growth over time if one were not so sure, one way or the other, of their possible survival; but with the Georgetown Round Table on Languages and Linguistics, it seemed to us instantly that this was going to be a successful venture. Of course, by 'successful venture' we could not have imagined that forty years later there were going to be hundreds of people in attendance, a conference lasting for days, and people coming from all over the world. Nevertheless, we did feel that it was definitely a success from the beginning.

Although I have to admit that I was not present at the founding of Georgetown University, I must admit that I was here at the beginnings of all the other events connected with languages and linguistics at whose anniversaries we are taking delight.

Naturally the one I remember most is the Center for Applied Linguistics. I remember that day in February, 1959, when there were three of us only

6 / Georgetown University Round Table on Languages and Linguistics 1989

there, and we sat looking at each other knowing that we were supposed to be a clearinghouse and informal coordinating body in the application of the findings of linguistic science to practical language problems. But we didn't know what we were supposed to do. We just sat around the table in that room, and there was no idea at that point where we were going to go, in what direction. I must say that I have been pleased that it has had fruitful years, and perhaps the thing that impressed me most was that a few years after it got started, it began to have imitators around the world. There are a lot of places now named 'Center for Applied Linguistics' so I think we made some kind of impact.

But I must say in particular here that I have been pleased through the years to see the close connection between the Center for Applied Linguistics and Georgetown University. Members of the staff at CAL have taught at Georgetown; Georgetown students have joined the staff at CAL; they have cooperated on many ventures together. And, indeed, I hope that continues.

I am very happy to join in the celebration of the occasion. And I would like, in particular, on my own behalf, to congratulate Georgetown University, the School of Languages and Linguistics, and the Georgetown University Round Table, and wish them many more years of successful growth. We will, to be sure, later in the evening, discover from Sir John whether all the growth has been truly progress or not, but from my perspective I have certainly enjoyed the association up to the present time and extend every good wish for the future. Thank you.

On behalf of the Georgetown Bicentennial

Charles L. Currie, S.J.
Director, Bicentennial Celebration

It is a real pleasure to welcome you to Georgetown in our Bicentennial Year. We are celebrating the many faces and dimensions of the University through some ninety programs, including thirty conferences. Among the most significant events on that calendar are the programs developed by the School of Languages and Linguistics. I want to publicly thank the faculty, students and staff of the School, and especially Dean Alatis, who from the very beginning has been most encouraging and supportive of everything we have tried to do.

The Bicentennial Intercultural Festival of Performing Arts is an extraordinary combination of fifteen productions in fifteen languages--almost entirely the work of our students and faculty. We've enjoyed seven productions so far, with eight to go. Enthusiastic audiences have enjoyed French, Chinese, Japanese, Urdu, Farsi, Arabic and Afro-American performances, and we look forward to Spanish, German, Russian, Hindi, Italian, Hebrew, Portuguese, and English performances in the near future. It has been an extraordinary experience to watch our students not only handle the languages well, but also perform with no little acting skill. Needless to say, these students gain a new appreciation for the language and culture they are studying. The Festival also brings the international community to campus and the campus to area embassies. The French Embassy was the site of our production of Claudel's *La Ville*. One memorable performance used Chinese, Japanese, Spanish and English--all in one production. The entire project has been under the direction of Dr. Roger Bensky of our French Department, assisted by Dr. Barbara Mujica of the Spanish Department, with substantial help from their colleagues in the other language departments.

The French Department has also been the focus for four programs commemorating the Bicentennial of the French Revolution: 'Communication and Media in Contemporary French Culture,' 'Images of America in Revolutionary France,' a symposium sponsored by the Montesquieu Foundation, and an 'International Congress on the History of the French Revolution.' The first two are directed by French Department faculty, and the latter two are being hosted by Georgetown.

In all of this we have been trying to express the best of Georgetown while focusing our vision and energy for an ambitious third century. One of the special delights of the Bicentennial Year has been to learn something of the fascinating history of Georgetown University and its people. If we think there are problems facing higher education today, imagine the problems which confronted John Carroll, the founder of Georgetown, trying to start his 'little

8 / Georgetown University Round Table on Languages and Linguistics 1989

academy' without a charter, a president, a faculty member, students, and, God forbid, money! Yet he moved ahead, and got the job done.

It is also interesting to recall that the first faculty member at Georgetown was brought over from France by Carroll to teach French and Latin, while he himself learned English from the President. The founders of schools in those days were quite extraordinary people. The entire faculty might consist of four or five people. Yale, in 1789, had three. These were men of great dreams and magnanimous vision. The job description that John Carroll wrote for the President of his tiny 'Academy on the Potomac' would rival the most demanding Search Committee's description today, when we joke that the qualifier is God, on a good day. But this was part of Carroll's great vision for Georgetown, to be the best and to attract the best. As we celebrate our Bicentennial, we like to think his vision and dreams are being lived today, always with the promise of doing better.

Thank you for being with us at this festive time.

Presentation of Bicentennial Medals to Henry and Renée Kahane

James E. Alatis

*Dean, School of Languages and Linguistics
Georgetown University*

As part of our celebration, Georgetown has struck a Bicentennial Medal to recognize those who have served this University, those whose friendships we value and those whose lives reflect the ideals of Georgetown. It has been my distinct pleasure to present medals to the Ambassador of France and the Mexican poet and Nobel laureate, Octavio Paz. And now, I shall do honor to the Bicentennial Medal and the University it represents by conferring this medal upon Renée and Henry Kahane for their innumerable contributions to the field of language scholarship and for the important part they have played in the education of so many of the leaders in our field today.

The man and woman whom we are honoring this evening with the Georgetown University Bicentennial Medal are an incomparable example of how bountiful a joint pursuit of academic excellence can actually be. This is truly an example of the whole which is even greater than the sum of its two eminent parts. After almost sixty years of incredibly productive scholarship and of harmonious partnership in marriage, Henry and Renée Kahane can be conservatively credited with a scholarly output of at least a dozen books and well over one hundred and fifty other outstanding publications dealing with various aspects of literary history and linguistics, such as etymology, Romance and Mediterranean lexicography, stylistics, morphology, dialectology, and child language.

In this year in which Georgetown University proudly celebrates its two hundredth year of existence, and together with its many alumni, we gather here to celebrate a milestone that few of us can hope to match--the fifty years during which this extraordinary couple has been teaching in the United States, first at the University of Southern California and, since 1941, at the University of Illinois, Champaign-Urbana--and we can pause to consider with awe the number of American students who since 1939 have had the good fortune to study under and be guided by them in this half century. Enriched and inspired by the unwavering legacy of their caring and enlightened humanism, many of these students have, in turn, become professors at various distinguished institutions of higher learning throughout the world. Indeed, it should not surprise us at all that even our School of Languages and Linguistics can boast at least four professors whose lives have been touched by the Kahanes and their teaching: Dr. William W. Cressey, Director of the Office of International Programs, Professors Peter H. Lowenberg and Richard Lutz of the

Linguistics Department, and Professor Roberto Severino of the Italian Department.

As for myself, I first met the Kahanes in 1955 when I went to Greece on a Fulbright scholarship. I was immediately impressed by them both, for their seemingly endless range of learning and scholarship as well as for their warm and unrestrained offer of guidance and friendship. I taught Greek to American Fulbrighters using their book *Spoken Greek* and I have used this text in my classes since the first day I came to Georgetown. An exemplary book, I must add, that has valiantly stood the test of time, and which was written according to the linguistic and pedagogical principles underlying the ACLS and the Linguistic Society of America's Spoken Language Series. When I was in the U.S. Office of Education, Henry helped me as an evaluator for research proposals in the area of linguistics and Romance languages. And later, he served also as an evaluator for the first school-wide self-study conducted by the SLL.

Their story is really a fascinating story. We are honoring Henry and Renée, and we are paying tribute to two unique and unusual scholars--the very best that our profession can ever hope to produce--who are linked to us and to our times in many ways.

Henry Kahane was born in 1902, the son of the Austrian-born Arthur Kahane, a writer on theatrical subjects and the literary adviser of the world-famous Austrian stage and cinema director Max Reinhardt. Richard Oswald, his mother's brother, was one of the earliest and most prolific pioneers of German cinema, and by the 1930s had directed about 150 films. From 1922-1930 Henry studied at the University of Berlin and received his doctorate in 1930 under Ernst Gamillscheg with a dissertation on 'The Designation of the Jaw in Galloromance.' In 1931 Henry Kahane married Renée Toole, a Greek student with not so Greek a name, and in 1932 was appointed Assistant in the Department of Romance Languages at the University of Berlin, receiving, in addition, a fellowship toward his '*habilitation*'.

Renée Toole Kahane was born in 1907, on the island of Cephalonia, Greece, and received her doctorate in 1931, also under Gamillscheg, with a dissertation entitled 'Studies on Word History: *toupin* and *bronze*.' Her father's family, of Irish descent, had settled on Cephalonia in the nineteenth century, when the island was under English domination. Her mother's family, Greek of Italian descent, had come to the island during the long period of Venetian rule. Especially on her mother's side Renée's family embodied many patriotic and humanistic virtues. Her maternal grandfather and great uncles had been leaders in the movement that brought the island group back to Greece, and were men of wide interests and great learning, amassing in their house overlooking the Ionian sea a uniquely rich library.

In 1933, after the Nazis' accession to power, the Kahanes emigrated to Florence, Italy, where they stayed until 1938, working partly at the University of Florence (Renée taught Modern Greek and Henry, at the University's Magistero, Romance Linguistics), and partly at a *Landschulheim*, a German school for refugee children which, because of the political situation, could boast a faculty body of the highest level. However, after the *Anschluss* of Austria in March 1938, the Kahanes were forced to escape once more, this time to Greece. From there they emigrated to the United States in 1939.

Their first joint publication in 1940, on the aptly ecumenical theme of Italian place names in Greece (*Italienische Ortsnamen in Griechenland*), marks thus the beginning of their extraordinary joint scholarly production and stimulating symbiosis, and of their crossing of many frontiers in pursuit of an intellectual Holy Grail and humanistic freedom and dignity, while unselfishly always holding the torch high so that others could see and follow on their path.

It is, therefore, with great pleasure that I bestow upon them Georgetown University's Bicentennial Medal and declare them son and daughter of Georgetown forever.

The last forty years: Real progress or not?

John Lyons

Trinity Hall, Cambridge University

1 Introduction. It is a great privilege to have been invited to give the opening address at this celebratory Georgetown University Round Table: celebratory of the bicentennial of Georgetown University, of the fortieth anniversary of what is currently the School of Languages and Linguistics, of the thirty-fifth anniversary of the famous 'Georgetown Experiment' in machine translation, and of the thirtieth anniversary of the Center for Applied Linguistics. It is a privilege of which I am fully appreciative, and I am grateful to the Dean of the School, Dr. Alatis, for having conferred this privilege upon me and to my good friend and former fellow student, Professor Francis P. Dinneen, S.J., for having persuaded him to do so--by what skillfully deployed jesuitical arts of rhetoric, casuistry, and moral blackmail, I know not. But I hope for his sake that, on this occasion, the end does indeed justify the means.

For the moment, however, I am more concerned with beginnings than ends. How to get started--whether I should even try to get started--was the biggest problem I had to resolve when I was invited to deliver this lecture. The brief that I was given in the letter of invitation was to deliver 'an historical lecture which would survey and evaluate the Round Table series and place them in the larger context of the developments in modern linguistics during these years', making appropriate reference to the 'many seminal papers in modern linguistics...given at the Round Table by prominent and influential linguists'. The phrase 'these years' was of course both anaphoric and deictic, and was itself appropriately contextualized. I will return to this point presently, as I will also return in due course to the phrase 'modern linguistics'.

My first reaction on receiving this invitation--when the warm glow of self-satisfaction induced by the flattering and persuasive phrase 'senior, established linguist' had subsided and some measure of common sense and self-knowledge had reasserted itself--was, I must confess, one of incredulity. The letter was quite clearly addressed to me, and internal evidence confirmed that I was the intended recipient. And yet the author knew full well, not only that, as he said, I 'was not intimately involved in the series' and might therefore, in principle, have 'the distance to effect the proper appraisal', but also that I had never myself been professionally concerned with language teaching, testing, and technology (which is the theme of this year's Round Table) and could hardly be expected to engage the interest of the experts who would be present on this occasion.

But then he also knew that I knew that he knew; and he had none the less invited me. Instead, therefore, of declining the invitation without more ado, tempting and flattering though it was, on grounds of incompetence, I had recourse to a bit of applied pragmatics. I reasoned myself, in true Gricean

style--or, perhaps better, in neo-Gricean style, since it required a pretty heavy use of the super-maxim of relevance--into a position of comprehending, and I trust not self-deluding, communicative cooperativeness. Not to put too fine a point on it, I decided that I understood what the Dean, being similarly rational and cooperative (it being of the nature of deans, anyway, to be such), would not have expected me to do what I am obviously not capable of doing, but had in mind, rather, something which I might more reasonably be presumed to be at least half-way competent at. And so I said yes. What I propose to do, under the rubric of my title, 'The last forty years: Real progress or not?', I will explain shortly.

Now, lest you should have started wondering whether I am myself at this point being irrelevant, lest you should have started asking yourselves, impatiently, when I am going to get started with the substance of my talk, I had better tell you now that I have done so, though somewhat deviously I admit, already. We are already *in mediis rebus*. I have been using terms and concepts very different from those that would have been used, or could have been used, at the first Round Table--*GURT-1*, as I shall call it, acronymically --in April 1950; and I have been soliciting, successfully I trust, your in-group complicity in my employment of them. Terms like 'deixis', 'anaphora', 'contextualization', 'applied pragmatics', and 'Gricean' evoke a very different paradigm, and a much broader or richer view of linguistics, than it was possible for anyone present at *GURT-1* to have been comfortable with.

Not even Charles Ferguson, the only person giving a paper this year who was also a speaker at *GURT-1*--and who for that reason deserves early and special mention, but also because he is someone whose own work, together with that of other contributors to the Round Table whom I shall mention presently, and of Henry and Renée Kahane whom we have just been honoring for the breadth and quality of their scholarship, has contributed significantly in the intervening years to the formation of the broader view of language with which we now all operate--not even he, or they, would have assumed, as casually as I have been able to do, a familiarity with concepts which, at that time, were either unknown to linguists or, if known, were treated with disdain and suspicion, as unredeemably unscientific, philosophical and--dare I say it?--traditional. Mainstream linguistics in general, and mainstream American linguistics in particular, was, at the beginning of the period with which we are primarily concerned, avowedly scientific, deliberately unphilosophical and assertively antitraditional. Things are very different now. And in the forty years of its existence--'these years', as they are referred to in my brief--the Round Table has played its own role in making them so. But more of this as we proceed. I have still not established what my brief calls the broader context, by supplying a referent for the phrase 'modern linguistics', on the one hand, and 'mainstream linguistics', on the other. To that task I now turn.

2 Chronology and numerology. Mention of the several anniversaries that we are celebrating on this occasion, coupled as it has been with a statement of my brief, 'to survey and evaluate the Round Table series and place them in the larger context of the developments in modern linguistics during these years', puts me in mind of Charles Hockett's *Presidential Address to the*

Linguistic Society of America on December 26, 1964, published, in *Language*, in Spring 1965 (Hockett 1965).

These dates are in themselves significant, falling as they do at the end of a short period, of perhaps five years, in which, as Hockett recognized, there had been a revolutionary change of paradigm in mainstream American linguistics. This was a change of paradigm to which *GURT* made its own notable contribution, especially--but not only--*GURT-15*, with its star-studded 'internationally formed Panels' deliberately set up (I quote from the editorial Foreword to the published *Report*) 'so as to permit unified efforts to reach consensual agreement and thus overcome the admittedly heterogeneous nature of contemporary linguistic outlook' (Stuart 1964: vii). For this year, I would remind you, we are celebrating, not only the several anniversaries listed on the program, but also the twenty-fifth anniversary of *GURT-15*, the first and so far the only fully international Round Table. But the whole period from 1961 to 1965 was of particular significance in the history of *GURT*, as also in the history of mainstream linguistics, American and non-American; they were vintage years for *GURT*; and this is evident from even the most cursory perusal of the only such special *GURT* volume to have appeared so far, *Selected Papers on Linguistics 1961-1965* (O'Brien 1968).

I am however running ahead of myself. Before I start placing *GURT* 'in the larger context of developments in modern linguistics during these years', I must, as I said, first establish a referent for the phrase 'modern linguistics'. Hockett's 1964 *Presidential Address* will help me to do just that. It will also set the chronological framework for the larger context and for my own commentary, dialogic, dialectical, and, at time, contrapuntal, on the significant events, or changes of paradigm, which Hockett also discusses.

Hockett's *Address* was remarkable in several respects, and I am referring to it for several different reasons. I did not myself hear it delivered, as some of you must have done, but I read it when it was first published; and I could well imagine the electrifying effect it must have had on both the Old Guard and the Young Turks (the latter term, by the way, is Hockett's and the former all but leaps out of the context, activated by the structuralist principle of paradigmatic contrast and made present, as it were, *in absentia*). If there are any of you who are not familiar with it, I would urge you to remedy that deficiency as soon as you can. And partly to encourage you to do so and partly for one of my own immediate purposes, I will now quote the opening two short paragraphs:

On 2 February 1786, in Calcutta, Sir William Jones delivered an address to the Asiatic Society, in which occurs a passage that has since repeatedly been hailed as the first clear statement of the fundamental assumption of the comparative method. We may justifiably take that event as the birth of modern linguistics. Between Sir William's address and the present Thirty-ninth Annual Meeting of the Linguistic Society of America there is a span of 178 years. Half of 178 is 89, a prime number. If we add that to 1786 we reach the year 1875, in which appeared Karl Verner's *Eine Ausnahme der ersten Lautverschiebung*. Thereafter, two successive steps of 41 years each--41 is also a prime number--bring us

first to the posthumous publication of Ferdinand de Saussure's *Cours de linguistique générale* and then to Noam Chomsky's *Syntactic Structures*.

I have allowed myself this bit of numerology because I know none of you will take it seriously. But behind the persiflage there is a sober intent. Our fraternity has accomplished a great deal in the short span of 178 years; yet, in my opinion, there have been only four major breakthroughs. All else that we have done relates to those four in one way or another (Hockett 1965: 185).

The trouble with numerological series of this kind is that, if they do not quite suit one's own purpose, they set one thinking, with no less sober intent, of alternative series which can replace or complement them. Noting, for example, that Hockett makes the birth of modern linguistics almost, but not quite, contemporaneous with the foundation of Georgetown University, I was tempted to find some equally significant date in 1789 (other than the date of the French Revolution of course), not as well known to linguists admittedly as 2 February 1786, and substitute that as a starting-point. Noting also that 1876 has equally good claims as 1875 for the status of *annus mirabilis* in the history of linguistics (and has been officially celebrated as such in a special centenary volume published by the Philological Society) and that Chomsky's 'Three models' paper antedated *Syntactic Structures* and appeared in 1956; noting too that poor old Rask, Bopp and Grimm do not get much of a look-in for their work in the 1820s; noting all this, I say, and not having the same commitment to the primacy of prime numbers. I was tempted to substitute 40 for 41 in Hockett's calculation of the length, in years, of the intervals between 'the major breakthroughs' that have been made in linguistics in the last two centuries. You will be relieved to hear that I have resisted the temptation. Although I have been invited to give a historical lecture, that is not the kind of slightly off-beat historical lecture that the Dean had in mind, I am sure, when he issued the invitation and that you were expecting to hear. And what is a year or two, anyway, between friends? I am sure that Hockett will not mind if I round off 203 to 200 and if I operate with 40 rather than 41, the more so as these round numbers conveniently split into decades and are also relatable more easily to quarter-centuries.

So, with that slight modification, I will accept Hockett's periodization of the history of the 200 years of what he calls modern linguistics. But I will now superimpose, upon the latter part of that period, contrapuntally as it were, a further two minor 40-year series, each of which will then be justified thematically in what follows.

Hockett, you will observe, was the fortieth President of the Linguistic Society of America. And he was commenting judiciously, perceptively, responsibly, and with a remarkable generosity of spirit, upon the demise of what had been, for the previous forty years, the dominant paradigm in mainstream American linguistics--a paradigm in the establishment and theoretical elaboration of which he had himself been prominent. In fact, his *Address* can be seen as giving formal recognition to the fact that at the end of this 40-year period there was no longer such a thing as mainstream American linguistics--American linguistics, as it had been presented in the

first edition of Joos's *Readings*, published (significantly) in 1957 and in American textbooks of the period. Mainstream American linguistics, self-confidently and triumphantly isolationist at the time of *GURT-1*, had been internationalized; and non-American, European, linguistics was well on the way to being Americanized. But more of my explanatory and justificatory exegesis in due course. I have already alluded to the role that *GURT* played in the change of paradigm that took place in mainstream American linguistics in the early 1960s; and I will fill in some of the details later; for this, as will be by now apparent, is going to be one of my major themes, in my discussion of the early years of *GURT*.

Superimposed upon Hockett's major periodization of linguistics in 40-year intervals since the time of the Neogrammarian revolution, we have, then, the first of my two contrapuntal periodizations, with 1925 and 1965 as the significant dates. And if you want to go back a further 40 years in the history of American linguistics, you will find that it was in the mid-1880s that Whitney published his *Life and Growth of Language*--one of the few works to have been explicitly acknowledged by Saussure in the *Cours* as having exerted a positive influence upon the formation of views that we now think of as characteristically modern, but also one of the first distinctively American works in the field of what we now call general linguistics. But I will not pursue the history of the American 40-year series any further back than 1925. This, you will recall, is the date of the foundation of the LSA and of the journal *Language*, the first issue of which contained, *inter alia*, the abstracts of Collitz's 'Scope and aims of linguistic science' (1925) and Boas's 'Problem of kinship among American Indian languages' (1925), both of which bear witness, in their different ways, to the Americanization of Neogrammarian linguistics, and the second volume of which contained Bloomfield's 'Postulates', heralding the development of that specifically American version of synchronic linguistics that was later called Bloomfieldian, neo-Bloomfieldian or post-Bloomfieldian, or descriptivism, whose progress is charted in Joos's *Readings* and whose demise, as I have said, is formally recorded by Hockett in his *Presidential Address*.

My second contrapuntal series of 40-year periods is, of course, the present *GURT* series. And I hope we can refer to it as such, as we celebrate the completion of the first 40 years and look forward confidently to the initiation and in due course the completion of the next 40-year period. I shall be making some predictions about the next 40 years, in my own modest attempt at a 'View toward the Future' in the final section of my talk. But to complete the numerology, by bringing in the quarter-centuries and in so doing establishing the chronology upon which I am pegging my history of both *GURT* and modern linguistics, let me add that its inauguration occurred just 25 years after the inauguration of the American series, of which it was in its first few years, fully representative--that is to say, it was inaugurated in the middle of the third decade of the current 40-year period of American mainstream linguistics, when descriptivism was at its height and at its most assertive.

Also, it is perhaps coincidental, but we may take it as symbolic, that the volume of *Language* which appeared in the year of *GURT-1*, volume 25, contained the obituary of Bloomfield and that the volume in which Hockett's

Presidential Address appeared in 1965 was the last of the 25 volumes edited by Bernard Bloch, his obituary appearing in volume 26. It follows too that the silver anniversary of *GURT* coincided with the golden anniversary of the LSA, both of which events were appropriately celebrated by the publication of commemorative, retrospective volumes and articles, upon which I have drawn and which I have included in the references.

Hockett, as we have seen, takes Sir William Jones's famous pronouncement of 1786 to mark the birth of the now 200-year-old period of modern linguistics. This is a defensible view. However, it is perhaps more usual to restrict the temporal scope of 'modern linguistics' to the period initiated by the publication of Saussure's *Cours* (1916); and this is what I will do. For what it is worth, the more or less conventional periodization of linguistics with which I shall be operating recognizes a two-way division by centuries and, within this, a five-way division by the publication dates of significant works, each of which we may think of as bringing about a change of paradigm, or, in Hockett's words, as constituting a 'major breakthrough'. Nineteenth-century linguistics is characterized by the predominance of comparative philology and of the historical, or diachronic, point of view; twentieth-century linguistics, by the development, triumph, and currently unchallenged supremacy of the opposing, synchronic, point of view.

As to the five-way division by dates, the five periods may be labelled for our purposes as follows (the labels are my own):

1. The age of the prophets: initiated by Jones (1786)
2. The age of the founding fathers: initiated by Rask (1818) or Grimm (1822)
3. The classical period (of comparative philology): initiated by the Neogrammarians in the mid-1870s
4. The post-Saussurean period (of modern linguistics): initiated by Saussure (1916)
5. The Chomskyan--and post-Chomskyan--period (of modern linguistics): initiated by Chomsky (1957)

It will be obvious from this schematic periodization of the subject that what we now think of as twentieth-century linguistics, or modern linguistics, did not start officially until well after the turn of the century and, conversely, that nineteenth-century linguistics has not only continued into the twentieth century, but is still practised by many specialists in more or less its classical mode. It is also hardly worth stressing that much of what we think of as modern linguistics can be found in earlier ages. Indeed, there is a sense in which later versions of modern linguistics, in particular Chomskyan generativism, have reasserted many of the aims of traditional, pre-nineteenth-century linguistics. This is a point that I will pick up later.

3 Mainstream linguistics and Whiggism. Hockett's periodization of linguistics, which I have adopted is, of course, a fairly conventional periodization of what he, and probably the majority of linguists, would regard as mainstream linguistics. The term 'mainstream linguistics', which I have been employing and shall continue to employ, is, however, historiographically

contentious. At this point, therefore, I had better say something to justify both my use of the term and my adoption of the viewpoint that its use implies, the more so, as the task that I have been invited to perform and the terms in which it was phrased oblige me to be properly historical.

The historiography of linguistics, as of science and scholarship in general, has what Henry Hoenigswald has recently referred to as its own 'Whiggishly naive (and not always innocent) apologetic linearities' (1986:176). The article from which I have taken this quotation is one of many works published recently that have sought to change the standard, Whiggish, view of the history of nineteenth-century and early-twentieth-century linguistics that is found in, say, Pedersen (1931) or Jespersen (1922) and has been repeated, more or less uncritically, in lectures and textbooks, including my own (Lyons 1968:21-52), down to the present day. It is incumbent upon me, therefore, to say something about this matter.

Clearly, if we are to have a proper sense of the history of our discipline we must free our minds, as far as we can, from the naive conception of the past as developing monotonically into the present along a single time-line; and as being periodizable, naturally and without a considerable degree of arbitrariness, into a definite number of segments with definite starting-points. We must recognize that, as a matter of historical fact, the course of progress—if progress it is—is much more complex and less determinate than this. We must also accept that such terms as 'modern linguistics' are just as much thematic, as chronological, labels. They should not be thought of as being applicable, therefore, equally and in the same sense, to all and only the work of theoreticians and practitioners of the discipline carried out or published in a particular chronological period. Specialized historiographical investigation has recently shown, moreover, that each of the great innovators or discoverers whose work we now see as seminal and epoch-making had his predecessors; and also that, in some cases, a particular scholar's own view of his aims and achievements would have been quite different from ours.

Linguistics is, of course, no different from other academic disciplines in this respect. But it is not the historical facts of the matter that are of primary concern to us in the present context. The so-called discovery of Sanskrit by Jones may not have been his discovery at all; and his famous words of 1786 may not have been intended to be understood prophetically or, in context, to have the emphasis that we have learned to give them. It may be that, not Grimm (1922), but Rask (1818), or someone else, should be credited with the discovery of what has entered into history as Grimm's Law; and whoever the discoverer was certainly did not think of it as a law in the classical, Neogrammarian, sense. The Neogrammarians, in turn, did not vanquish all comers, immediately and decisively, in either 1875 or 1876.

As for Saussure, whom I have taken as the eponymous founder of modern, post-Saussurean, linguistics, he was by no means the first person to think of languages as two-dimensional relational systems, or to distinguish the synchronic from the diachronic point of view; and his posthumous *Cours* (1916), far from changing the whole character of linguistics overnight, took some decades before it came to be recognized, retrospectively, as the European structuralist's bible. In this connection it is worth noting that, although Hockett, in his LSA *Presidential Address*, chose the publication of

the *Cours* as one of his historically pivotal events, he had not mentioned it in his own textbook, published just a few years earlier (1958). Bloomfield had reviewed it in 1924, but he too had not listed it in the otherwise comprehensive bibliography of his 1933 textbook, the American structuralist's bible; and his review did not appear in the first volume of *Language*, as one might have expected it to have done, if the *Cours* had been recognized as a truly seminal work at that time. It is not cited in any of the textbooks or surveys of American structuralism of the 1940s and 1950s, such as Bloch and Trager (1942), Carroll (1953), the first edition of Gleason (1955), Hall (1950), Nida (1946), or Sturtevant (1949). Indeed, it was not mentioned, to the best of my knowledge, in any American textbook until we get to the second edition of Gleason (1961). Rulon Wells's review, included in Joos's *Readings* (out of chronological sequence) before Bloomfield's (1926) 'Postulates', did not appear until 1947; and it was published, significantly, not in *Language*, but in *Word*, the New York-based journal of the European, or non-American, currents in postwar American linguistics. The *Cours* was eventually reviewed in *Language*--habilitated, rather than rehabilitated, as it were--in 1972; but then it was Engler's critical edition, not the original, that was reviewed; and Koerner's review, quite properly, was historiographical, rather than historical. I have already emphasized that the distinction between American and non-American mainstream linguistics, real enough during the first 25 years of the LSA's existence and the first decade of *GURT*, eventually disappeared in the first half of the 1960s. What I have just said is but part of the evidence for this statement.

Hockett's (1965) retrospective adoption of Saussure as one of the founders of structural linguistics, American as well as European, can be criticized, then, as being historically incorrect--in one understanding of 'historical'. But there are contexts--the context in which he was speaking in 1965 is one of them and the context in which I am speaking today is another--in which the retrospective, quasi-mythological and ideological, evaluation, and reevaluation, of the past is more to the point than setting straight the historical record; there are contexts in which the mythological dimension, and the retrospective readjustment of history, is itself a matter of historical fact and at times a historic event. The mythological Saussure is now of far greater historical importance than the so-called historical Saussure. And so it is with the other adoptive ancestors and eponymous heroes, including Bloomfield and Chomsky, of what we call, Whiggishly (and, in so far as we are conscious of battles yet to be won, perhaps not always innocently), mainstream linguistics. Those who are interested in following up some of the recent work which does aim to set straight the historical record may consult, as I have done, some of the references that I have appended to this paper, especially the admirable treatment in Hymes and Fought (1975).

Having protested my cognizance of the complexity of the actual historical record and my, not necessarily innocent, reasons for proceeding, regardless, to operate with the Whig and by now conventional, quasi-mythological, periodization of so-called mainstream linguistics, I will proceed to the second part of my paper, which divides into four short sections, each of which is simultaneously or concurrently chronological, retrospective and thematic, and a conclusion, which is both resumptive and prospective. My

treatment of the themes that I have selected for discussion will be of necessity broad-brush, brief, and personal, but not, I trust, irresponsibly so. It will not be possible to make detailed references to, or even to mention, more than a small number of the contributions that have been made to *GURT* over the years. I hope, however, that the broad generalizations that I make will be apposite and that my mention of particular scholars by name, always *honoris causa*, will not be judged invidious.

I now turn to the first of my chronotheematic periods, if I may so describe them.

4 American and European structuralism. The relation between American (Bloomfieldian, neo-Bloomfieldian or post-Bloomfieldian) and European (post-Saussurean) structuralism has been frequently and intensively discussed, from both a chronological and a thematic point of view. Not surprisingly, there is far from being universal agreement on all details. Let me begin therefore by stating, very briefly, my own view of what they have in common and what divides them--or used to divide them before they merged in the 1960s.

What united American and European structuralists in the period in which *GURT* was inaugurated, and what still unites them today, was the view that a language is a system, or structure, in which the elements or units (phonemes, morphemes, words, monemes, or whatever) derive, not just their essence, but their very existence from the relations of functional contrast, or equivalence, that they contract with one another in the system. That may not be how all self-declared structuralists would express themselves (I have always favored a Hjelmslevian version of post-Saussurean structuralism)--there might be some dispute, for example, about inserting 'existence' as well as 'essence', if only because this lexical opposition is too redolent of Aristotelian scholasticism--but I think it is fair to say that all versions of structuralism, American or European, are in fact expressible in these terms, with differences of emphasis or terminology at particular points. I think it is also fair to say, if I may quote my first mentor in modern linguistics, W. Sidney Allen, that by the time of *GURT-1*, it was as anachronistic to refer to structural linguistics, both in Europe and also in America, as it was to refer to the talking pictures, or the talking movies, as the case may be (Allen 1957). And that, incidentally, is still true. We are all structuralists now. Structuralism, in the broad-church sense of the term, as I have just defined it, is by no means incompatible with generative grammar and generativism, which will be the subject of my second chrono-thematic section.

What characterized American structural linguistics--more precisely, post-Bloomfieldian American structural linguistics--and distinguished it from various schools of European structural linguistics included, *inter alia*, a commitment to a certain kind of scientific physicalism; an aversion to the so-called mixing of levels (i.e. the use of grammatical prerequisites in phonology and of semantic considerations in either phonology or grammar); a tendency to formulate the principles of linguistic theory in terms of descriptive method and to formulate methodology in terms of the specification of what Chomsky later called discovery procedures; and a refusal to accept that meaning was a fit subject for serious and systematic investigation (in the current state of

linguistic theory and practice) and was not as fully a part of language as phonology and grammar. Once again, particular details of this all too brief characterization of American structuralism are open to challenge; and I would be the first to admit that it requires considerable expansion and qualification to make it historically satisfactory. But it is not, I submit, a travesty or caricature of what it attempts to portray in a few swift brush-strokes. And no one would deny, I think, that it is at least mythologically accurate.

By the end of the 1950s, when the change of paradigm commonly referred to as the Chomskyan revolution, was already well in train, though it is only retrospectively that we can see that this was so, the differences between American and European structuralism were also well on the way to being eliminated. As I said earlier, the former was being internationalized and the latter Europeanized; and not only *GURT-15*, but several of the preceding sessions of the Round Table, played their full part in this process. The process was accelerated particularly, no doubt, by the Ninth International Congress of Linguists held at Harvard in 1962. This provided its own forum for interschool discussion with the widely circulated *Preprints* and, in due course, the voluminous and highly influential *Proceedings* (Lunt 1964). It also furnished the occasion for the publication of a large number of articles, in journals and special volumes, devoted to the presentation of this or that school's doctrine and contribution to the mainstream. By the mid-1960s, when Hockett delivered his *Presidential Address*, it was becoming less and less sensible, in my view, not only to draw global distinctions, of the kind that I have drawn, between American and European structuralism, but also to distinguish sharply between one school of European structuralism and another. Autobiographical evidence certainly suggests that this is so.

And if I cite what I have called autobiographical evidence, now and subsequently, I do not do so, as you will see, gratuitously or vaingloriously. When my own views on linguistics were formed in the late 1950s and the very early 1960s, 'I had the inestimable good fortune', as I said in the Preface to my *Introduction to Theoretical Linguistics* (1968), 'to come under the personal supervision of four inspiring teachers: W. Sidney Allen, C. E. Bazell, F. W. Householder, and R.H. Robins'. All four of them were intimately familiar with several schools of structuralism, American and European, and they each had their own particular, responsibly eclectic and well-considered, approach to the major theoretical issues of the day. Sidney Allen I have already quoted. As to the other three, each of them has a connection, direct or indirect, with the Institute of Languages and Linguistics at Georgetown and with the Round Table. Robins, who participated in *GURT-25* (1974), was not only my teacher, but also Frank Dinneen's, and his influence shows, not only in my textbook but also in Frank's, published in the previous year and based on the courses that he had been giving for some time to 'undergraduate majors in both linguistics and languages' at Georgetown (1967:vi). Bazell was also one of Frank's teachers, in London, in the late 1950s; and Bazell's contribution to *GURT-15*, 'Three misconceptions of grammaticalness', though it was probably not properly understood by more than a minority of linguists who heard it delivered or read it shortly afterwards in its published form, looked forward, as we can now see in retrospect, to a time when the distinctions that he was drawing would be taken for granted and when it would have become obvious

that the failure to draw such distinctions had confused, hopelessly, not only the boundary between syntax and semantics, but also the boundary between semantics and pragmatics, on the one hand, and between linguistics and ontology, on the other.

But it is Fred Householder, and his *GURT-7* paper, 'Rough justice in linguistics' (1957), that I wish to refer to more particularly here. It was Householder who hired me in 1960 to work for a year, and again in 1963 to work for a further six months, on his project for the syntactic analysis of English by computer, supported by the U.S. Air Force. It was during that time, therefore, that I became familiar with the research on machine translation being done at Georgetown, Rand, Ohio State, and other centers. It was also during this time, working on the project at Indiana University and teaching one or two courses in linguistics from the standard American textbooks of the time, attending seminars organized by Fred Householder, Thomas A. Sebeok and Charles F. Voegelin, that I filled in some of the many gaps in my knowledge of American structural linguistics, upon which I have been drawing for my commentary today. And it is perhaps not surprising that my views on the topic that I am at present discussing should be so similar to Householder's. Householder was charged, in 1956, with much the same task as I have been charged with today: to comment upon the current state of linguistics, with particular reference to the Round Table series, and to make some predictions as to probable future developments. The title of his paper, you will recall, was intended to point to a *via media* between the 'God's truth' and the 'hocus-pocus' approaches to linguistics; and these terms, which were widely used in the early 1950s, were drawn, of course, from Householder's famous review of Harris's *Methods*, the very culmination or pinnacle of American structuralism--the work, too, on which Chomsky, as it were, cut his linguistic eye-teeth.

I have referred to Householder and to his *GURT-7* paper for three reasons. First, it allows me to make public acknowledgement of my indebtedness to him, on the most appropriate occasion that I have ever had in the United States, for bringing me here in 1960 and for giving me the opportunity of improving my background in linguistics. Second, it establishes my credentials, I hope, as someone who, though coming from another tradition, can speak from first-hand knowledge about American linguistics (and, as it happens, machine translation) in the early 1960s. Third, it saves me the trouble, and you the time spent in listening to me, of saying any more about the theoretical issues that were debated in the early years of *GURT*. Everything that I might say was said, and in much the same way that I would wish to say, in Householder's 'Rough justice' paper.

But I should perhaps highlight the contact that the Round Table has established from the very outset with leading theoreticians and practitioners by mentioning a few more names. Hoenigswald, from whom I borrowed my quotation about Whiggism and the history of linguistics, was President of the LSA in 1958, as Hockett was in 1964. He was also, as many of you will know, yet another of the many distinguished American linguists who answered the call from Georgetown in 1950 and on many subsequent occasions. (In so expressing it, but spelling the word *call* with a distinctively lower-case initial *c*, I am deliberately evoking that other Call, spelt with a capital *C*, which led

to the foundation of the Linguistic Society of America in 1925 and of the journal *Language*, which I have already identified as one of the most significant events in the history of mainstream American linguistics.)

Hoenigswald, like Charles Ferguson and others I am about to mention by name, participated as a speaker in *GURT-1* and in several other Round Tables in the years that followed. His contribution was on the teaching of Latin. Let me exploit that fact, therefore, to make the following comment (without however wishing you to read too much of the Vergilian context into my paraphrase): *haud minima pars fuit* 'he was himself no small part of the proceedings'. The comment applies, not only to Hoenigswald, but also to all the other distinguished scholars who answered the *GURT* call on that or subsequent occasions, including notably, in the earlier period, Henry Lee Smith, George Trager, Archibald Hill, Martin Joos, Bernard Bloch, Paul Garvin, Freeman Twaddell, Morris Halle, Einar Haugen, Norman McQuown, Robert Stockwell, William Bright, Joseph Greenberg, Henry Gleason, Ward Goodenough, Carleton Hodge, Winfred Lehmann, William Austin, Leon Dostert, Floyd Lounsbury. I could extend the list; but I have taken only the first decade of Tables, for the moment; and I have been perforce selective. My selection, it will be obvious, is far from being unmotivated. These are all names to conjure with. My comment also applies, not only to their participation in *GURT*, but has wider scope. Each of the scholars that I have named, and others that I will refer to in due course, has played a major role, not only here, from time to time at Georgetown, giving one of the 'many seminal papers in modern linguistics' delivered at the Round Table 'by prominent and influential linguists', to which Dean Alatis referred in his letter of invitation and which he invited me to 'place in the larger context', but also on the wider, national and international, scene. Most of them published frequently in *Language* and wrote widely quoted books and articles which were taken to be typical of the best of American linguistics in the 1950s; several are represented in Joos's *Readings*; each of them, in short, played an important part in the formation and development of what I am calling, Whiggishly, mainstream American linguistics. Their participation in *GURT* is, therefore, indicative of the relation that the Round Table established from its earliest days with leading American theoreticians and practitioners. As further hard, statistical evidence of the strength of the connection established and maintained with the most eminent American linguists of the day, I may be permitted to mention (as the fruits of my research) the fact that no fewer than three-quarters of the presidents of the LSA between 1949 and 1988 have participated, at least once, in *GURT* during that time. The list includes not only Charles Ferguson, Henry Hoenigswald, and Charles Hockett, but also Winfred Lehmann, and, of course, Henry Kahane, who are present today. The determination of the exact number and of the other names I will leave as an exercise for the reader.

5 The Chomskyan revolution: Generative grammar and generativism. That there was a revolutionary change of paradigm in mainstream linguistics, initiated by the publication of Chomsky's *Syntactic Structures* in 1957, is widely accepted (though not, of course, undisputed). What exactly was revolutionary about it; whether it really was, as Hockett said, one of 'only four major breakthroughs' in the advance of linguistics from its primitive, prescientific,

beginnings; just how much of what is conventionally attributed to Chomsky was truly original to him: these are all questions about which there is still plenty of scope for legitimate disagreement. But they are questions that cannot, and need not, be discussed here (see References).

There is just one of the many aspects of the so-called Chomskyan revolution that I have chosen for brief commentary in the present context; and it has to be, of necessity, even briefer than, ideally, it ought to be. This has to do with the difference between generative grammar, properly so called, and what may be referred to, contrastively, as generativism. A generative grammar, as we all know, is a mathematically precise, formalized, system of rules which, operating upon or in association with a vocabulary, characterizes the sentences of a language (the language that it generates) by defining them to be well-formed in that language and assigning to each of them one or more structural descriptions, and perhaps also, though not necessarily, one or more semantic interpretations. And the theory of generative grammar is that branch of theoretical linguistics which deals with the formalization and associated topics: such as the hierarchization of grammars in terms of expressive power, problems of weak and strong equivalence, and so on.

It is my view, and it was Hockett's (1965), that Chomsky's work on the theory of generative grammar and his development of particular so-called models of generative grammar constituted a 'major breakthrough', which, as far as we can be sure of these things, has taken us along the path of progress. There can be no going back. This does not mean, of course, that all or any of the details of Chomsky's formalization are inviolate. Many details have indeed been changed, several times, both by Chomsky and by proponents of alternative versions of more or less Chomskyan or non-Chomskyan generative grammar in the last thirty years. What it means is that we are henceforth committed irrevocably to the mathematicization of linguistic theory, one of whose consequences is the conversion of linguistic theory into theoretical linguistics (see section 7).

By generativism, in contrast with the theory of generative grammar, I mean a set of philosophical, psychological, and metatheoretical assumptions and doctrines which are logically unconnected with generative grammar as such and, for the most part, with one another. What these assumptions and doctrines are is made sufficiently clear perhaps by mentioning the terms 'universal grammar', 'linguistic competence', 'the species-specificity of language', etc., and asking you to give them the interpretation that Chomsky has, for the last quarter-century or so, been associating with them.

By far the most controversial part of generativism, no doubt, is its association with mentalism and its reassertion of a particular version of the traditional philosophical doctrine of innate knowledge. Most of the other doctrines and assumptions that characterize it it shares with post-Bloomfieldian structuralism, out of which it emerged, or with other schools of European, post-Saussurean, structuralism, with which, in various ways, it has now associated itself. I do not wish to take a stand on the validity of any or all of the cluster of doctrines and assumptions that I am labelling, as a whole, as generativism. I simply wish to emphasize the importance of not confusing generative grammar, properly so called, with generativism, as I have described

it here. This is all the more important, in that Chomsky himself uses the term 'generative grammar', as do his followers, to refer to both.

It is worth remembering, however that there was little or no trace of generativism in *Syntactic Structures*. In fact, apart from his attack on discovery procedures, Chomsky's (1957) presentation of transformational-generative grammar is fully compatible with post-Bloomfieldian methodological and metatheoretical principles. In this connection, it is interesting to read Chomsky's *GURT-8* (1955) paper, 'Semantic considerations in grammar', which formed the basis for Chapter 9 of *Syntactic Structures*, and the panel discussion which ensued. Generativism did not make its appearance, in Chomsky's own publications, until he wrote his long paper for the Ninth International Congress of Linguists. Since then, however, he has, of course, been a consistent advocate of his particular version of universal grammar, as a set of genetically transmitted principles which determine the acquisition of language, and has insisted that his own interest in generative grammar is motivated by his wish to give the best formal account of this. For most people, and certainly for philosophers and psychologists, it is Chomsky's powerful advocacy of generativism, rather than his work on the theory of generative grammar, that is at the heart of the so-called Chomskyan revolution. For my own part, without wishing to deny the interest or importance of this aspect of the Chomskyan revolution, I am inclined to agree with Hockett (1965), for whom it was the former that will prove to be of permanent significance. I also agree with Hockett, that, far from being an antithetical reaction to structuralism, Chomsky's mathematicization of linguistic theory and description can be seen as the product of the post-Bloomfieldian striving for accountability and rigor. And Chomsky himself acknowledged this to be so in his paper for the Ninth International Congress of Linguists.

Having said this, however, I, in turn, must also acknowledge that it was Chomsky's formulation, and robust defense, of the several logically distinct tenets of generativism that has determined the whole course of mainstream linguistics since the mid-1960s and has been the major catalyst in the broadening of the field of linguistics to which I referred in my opening remarks.

I cannot of course mention more than a small fraction of the 1960s *GURT* participants or comment upon their individual contributions. But I may perhaps add a few more names to the list of those I have already mentioned (some of which figure prominently throughout the second, as well as the first decade of our 40-year period): Robert Lado, André Martinet, Kenneth Pike, Michael Halliday, Sidney Lamb, William Labov, Charles Fillmore, Paul Postal, Emmon Bach, Frank Palmer, Terence Langendoen, Joshua Fishman, John Gumperz, Walter Cook, Raven McDavid, Charles Kreidler, Eric Hamp, John Carroll, John Lotz, Robert Longacre, Peter Strevens, Peter Ladefoged, Francis Nelson, and Seymour Chatman. That is an impressive list, impressive not only for the quality of the contributions made by the scholars I have listed, but also for their theoretical diversity. Anyone who wishes to get some idea of what it was like to be involved in the Chomskyan revolution in the 1960s could do worse than read not only the *GURT* papers of this period, but also the panel discussions.

6 Semantics and pragmatics. My third theme is one upon which I have written at inordinate length elsewhere. On this occasion, I will be brief, if only to confound my critics, who have frequently accused me, with some justice perhaps, of never using one word when five will do.

I have put 'Semantics and pragmatics' in the third place because the 1970s, the third decade of the present *GURT* series, were the years during which semantics came into its own as a fully enfranchised constituent of modern linguistics and the distinction between semantics and pragmatics gradually came to be widely, if not universally, accepted by linguists. And *GURT*, as we shall see, has played a full part in that process.

A thumbnail history of mainstream linguistic semantics in the 40-year period with which we are retrospectively concerned would note three phases, or aspects, of its development: (1) what I have referred to as its enfranchisement; (2) a broadening of the field to cover nonlexical meaning and to integrate the coverage of this with lexical semantics and also with syntax in terms of the principle of compositionality; (3) the elaboration, for the first time, of at least the outlines of what many would take to be a viable and properly formalized theory of linguistic semantics.

Most linguists nowadays are at least familiar with the terminological distinction between 'semantics' and 'pragmatics'. This would not have been the case until, let us say, the late 1970s. Although the distinction is now familiar enough, how it is drawn varies from author to author, according to their definition of 'meaning', on the one hand, and of 'language', on the other. Here is a list of propositions, by no means complete, specifying various criteria which have been invoked in this connection (cf. Lyons 1987:157):

- (1) Semantics has to do with meaning, and pragmatics with use.
- (2) Semantics has to do with competence, and pragmatics with performance.
- (3) Semantics has to do with the conventional, and pragmatics with the nonconventional, aspects of meaning.
- (4) Semantics is a matter of rules, and pragmatics of tendencies, principles, maxims or strategies.
- (5) Semantics deals with truth-conditional, and pragmatics with non-truth-conditional, meaning.
- (6) Semantics deals with literal, and pragmatics with nonliteral, meaning.
- (7) Semantics has to do with the meaning of sentences, and pragmatics with the meaning of utterances.
- (8) Semantics deals with context-independent, and pragmatics with context-dependent, meaning.

The effect of applying the distinction between semantics and pragmatics, in terms of one or more of these criteria, is to narrow the field of what has been traditionally, if only pretheoretically, thought of as semantics by classifying some kinds of meaning as pragmatic, rather than semantic.

It is readily demonstrable that several of the criteria that I have listed here are in conflict. For example, there are conventionalized non-truth-conditional aspects of both lexical and nonlexical meaning, and they are just as much a matter of competence as is a knowledge of the truth-conditional

meaning of words and constructions; it is not only the meaning of non-declarative sentences (interrogative, imperative, etc.) that causes problems for truth-conditional semantics, but also the meaning of many, if not all, declarative sentences; literal meaning must be distinguished both from truth-conditional, or propositional, meaning and from conventional meaning; and so on.

As I have said, I have written extensively on these topics elsewhere. Personally, I prefer to employ the term 'semantics' in its traditional, broader sense: that is to say, the sense in which it has been used by linguists until recently. But this is, of course, simply a question of terminological preference. I do accept that it is desirable to recognize within the broader field a narrower subfield, microlinguistic semantics, and, not just one, but several intersecting subfields of macrolinguistic semantics, of which pragmatics, under a particular definition, might be one. For what it is worth, I think that the best way of drawing the distinction between microlinguistic semantics, semantics in the narrow sense if you will, and pragmatics is on the basis of a properly elaborated distinction between sentence-meaning and utterance-meaning and this, in turn, involves, not only distinguishing sentences from utterances, but also propositions from both sentences and statements, on the one hand, and from propositional content, on the other; reference from denotation; utterances, as locutionary and illocutionary acts, from utterances as the products of those acts; and so on.

Further progress in semantics-cum-pragmatics, I would say, depends upon a more general acceptance of these truths, which, though blindingly obvious to me, are curiously less so to others. It also depends, in my view (my 'view toward the future'), upon a more general recognition of the importance of locutionary and illocutionary subjectivity, even in microlinguistic semantics, and of the fact that, contrary to what is commonly supposed, natural languages differ in their expressive power, even truth-conditionally, and are, for that reason, not intertranslatable. But that is currently my King Charles's head, and there I must stop.

It remains only to note that, although the Round Table's interest in semantics and what is now called pragmatics goes back way beyond the 1970s (one thinks, for example, of *GURT-11, 12* and *18*), what I have selected as my theme for this third decade is well represented, especially in *GURT-27* (1976), where, in addition to others already mentioned, Ray Jackendoff, Geoffrey Leech, Gilbert Harman, and Barbara Hall Partee were among the participants.

But the 1970s were also the years when what I shall call macrolinguistics, in particular sociolinguistics, came to be more highly developed. And both Georgetown University and *GURT* were strongly represented in its development. It suffices to mention from the home team Roger Shuy, Ralph Fasold, and, if we look forward to the 1980s, Deborah Schiffrin and Deborah Tannen, and the Round Tables that they organized in pragmatics-cum-sociolinguistics. I have not been able to discuss this aspect of the question, but the boundary between pragmatics and sociolinguistics is, of course, a matter of dispute. I would have you bear this in mind, as you consider what I have to say in the next section. Although the boundaries between (autonomous) microlinguistics and macrolinguistics are reasonably sharp in respect of phonology,

morphology, and syntax, they are less so when it comes to semantics-cum-pragmatics. The post-Bloomfieldian attitude to meaning was not wholly unmotivated.

7 Theoretical and applied linguistics. My next and final theme has to do with the distinction between theoretical and applied linguistics; and I will treat it once again from both a historical, or retrospective, and also from a prospective, not to say prescriptive, point of view. It is a theme that is perhaps, inherently, less obviously related to a chronological period than my other three themes have been. But, I think it is fair to say that the points that I am making in connection with this distinction have only become clear, at least to me, in the fourth decade of the first 40-year period of *GURT*, which has now just ended. And it is a personal, indeed a rather idiosyncratic, view of the matter that I am giving.

This theme is obviously highly relevant to the evaluation of the last 40 years of *GURT* and to the formation of 'a view toward the future'. Several of the *GURT* volumes have included the term 'applied linguistics' in their title or in the rubric for one of the panels, and many of the others have been devoted, in whole or in part, to topics such as language teaching or machine translation that are commonly held to fall within the scope of the term. But there have also been Round Tables that have dealt, wholly or partly, with different approaches to linguistic theory (generativist, stratification, tagmemic, etc.) or to the relation between theory and application. One thinks especially, in this connection, of *GURT-10*, *15*, *18*, *22*, and in the period that I am now discussing, *36*, *37*, and *38*. Indeed, it can be fairly said, I think, that one of the most valuable services that *GURT* has rendered to linguistics as a whole in the last 40 years has been to keep this evolving and dialectical relation under review and to provide a forum for its constructive development.

It would be presumptuous and foolish of me to attempt to say anything about applied linguistics in this company. The cobbler should stick to his last. I will therefore concentrate upon the theoretical side of the relation. I would emphasize, however, that it is a dialectical relation with feedback in both directions. There is nothing so practical as a good theory (properly applied); and theory without the empirical controls and insights that come from application is not likely to win friends and influence people for very long.

Although I shall say nothing of substance about applied linguistics as such, I should perhaps explain how I am using the term. I do not need to tell anyone here that the term 'applied linguistics' has been employed, confusingly at times, in several senses. Looked at from the point of view that I am adopting, it has been used: (1) with reference to the application of any branch of linguistics to any practical purpose; (2) with reference to the application of any branch of linguistics to a restricted set of purposes (especially language teaching and machine-translation); (3) the application of some, but not all, branches of linguistics to any practical purpose; (4) the application of some, but not all, branches of linguistics to a restricted set of practical purposes. And pedagogically, of course, the term has frequently been used to include the application of the relevant parts of disciplines other than linguistics (psychology, computer science, etc.) to particular practical purposes.

When my *Introduction to Theoretical Linguistics* (1968) was published, over twenty years ago, it was hailed by Bar-Hillel as 'The first [book of its kind]... to carry the long overdue adjective, "theoretical" in its title' (1969: 449). Another review of my book, more critical than Bar-Hillel's, was Starosta's, published in *Language* (1971). This rightly drew attention to my failure to develop, seriously and consistently, the implications of the programmatic opening sentence: 'Linguistics may be defined as the scientific study of language'. This criticism was, I think, well founded. And I would now concede, further, that I not only failed to define 'theoretical linguistics' (tacitly identifying it with 'general linguistics' or even with 'linguistics' *tout court*), but also adopted far too narrow a view of its subject matter. In effect, I restricted the scope of theoretical linguistics to what I would now characterize as general, theoretical, synchronic microlinguistics (cf. Lyons 1981a: 34-37). I still think that this constitutes the central and most distinctive part of theoretical linguistics. But I certainly do not believe that diachronic linguistics is intrinsically less theoretical than synchronic; that such branches of macro-linguistics, sociolinguistics, psycholinguistics, or stylistics are, by virtue of their data and the questions they address, less theoretical than microlinguistics; or even (though I grant that this is more debatable) that descriptive linguistics (i.e. the description of particular language-systems) is necessarily less theoretical than general linguistics (i.e. the study of language in general).

The point I am making here is that the field of linguistics can be structured in terms of at least four logically distinct dichotomies: general vs. descriptive, synchronic vs. diachronic; micro- vs. macro-; and theoretical vs. applied. And at this point I welcome the opportunity of acknowledging that I have borrowed the terminological distinction between micro- and macro-linguistics from George Trager's (1949) programmatic distinction of 'the field of linguistics'. George Trager was a frequent participant in *GURT* for many years and a doughty champion of old-style structuralism. But I too, as will be evident, am also an unregenerate structuralist, after my fashion.

The view of theoretical linguistics that I put forward in my 1968 textbook was, therefore, more restricted than I now think it ought to have been. Furthermore, linguistic theory has moved on in the last twenty years. There is now, as there was not in 1968, a well-established field of research and scholarship which is referred to by those who work in it as 'theoretical linguistics'. This has close links with logic and the philosophy of language; its hallmark is formalization; and it embraces, not only natural languages, but at least that (heterogeneous) subclass of nonnatural languages whose investigation is held to be revealing in the study of natural languages. The developments to which I have just referred make my 1968 use of the term 'theoretical linguistics' seem very dated. And, if I were writing such a book today, I would prefer 'linguistic theory' to 'theoretical linguistics'. But that is another story (Lyons 1990).

There are just three points that I want to make on the basis of what I have said so far. The first has to do with the distinction between theoretical linguistics and general linguistics, which, until recently, could reasonably be conflated (as indeed I conflated them in my 1968 *Introduction*). One of the consequences of the Chomskyan revolution has been the emergence of theoretical descriptive linguistics, distinguishable, on the one hand, from theoretical

general linguistics and, on the other, from nontheoretical, including applied, descriptive linguistics. (I will come back to the distinction between 'non-theoretical' and 'applied' in a moment.)

Theoretical general linguistics is that branch of the subject which sets out to provide a formally precise, intellectually satisfying, answer to the central defining question of linguistics, 'What is language?', understood, however, in a particular way. The question 'What is language?' can be addressed from several points of view and can be answered in several different, but equally legitimate, ways according to the point of view that is adopted. Theoretical linguistics, founded upon the Saussurean and post-Saussurean trichotomy of *langage*, *langue* and *parole*, interprets the question as meaning 'What is a language?'. And the different branches of theoretical linguistics adopt characteristically different points of view and consequently postulate different kinds of theoretical constructs in the answers they give to the question: they postulate different models of the language-system (*la langue*). Theoretical microlinguistics (often called 'autonomous linguistics' and associated historically with both European and American structuralism) adopts the point of view expressed by Saussure, or rather his editors, in the famous final sentence of the *Cours*: 'la linguistique a pour unique et véritable objet la langue envisagée en elle-même et pour elle-même' (1916: 317). It is the controversial 'en elle-même et pour elle-même' which distinguishes theoretical microlinguistics from the various subbranches of theoretical macrolinguistics: psycholinguistics, sociolinguistics, etc. They too have their own distinctive conceptions of the language-system (*la langue*). It must not be thought that sociolinguistics or psycholinguistics, or the other branches of macrolinguistics, can dispense with the distinction between the system and the process (or its products). This will be my second point. But before that I must make explicit something else which is not immediately obvious.

This is the fact that *la langue* in the famous passage from the *Cours* that I have just quoted (and which is frequently mistranslated into English as: 'Linguistics has as its unique and true object language [it should be "the language", i.e. "the language-system": "language", without the article, would translate *le langage*] considered in itself and for itself') can be interpreted either generically or specifically. Its generic interpretation defines the field of theoretical general linguistics; its specific interpretation, that of theoretical descriptive linguistics. A Chomskyan generative grammar is a theoretical description of a language (i.e. a language-system), namely the language that is generates. A generative grammar of any language, in short, is a theory of that language: more specifically, a theory of the well-formedness of the sentences of the language.

My second point arises, then, directly from the distinction that I have just drawn between theoretical microlinguistics and theoretical macrolinguistics, and nowadays it hardly needs to be emphasized. It is that the distinction between microlinguistics and macrolinguistics is independent of the distinction between theoretical and applied linguistics. There is, in principle, a theoretical aspect to every branch of macrolinguistics. If this was less obvious in the past than it is now, with a better understanding of the differences between the theoretical and the nontheoretical, this is simply because in such areas of applied linguistics as language-teaching it has always been essential to take the

broader, rather than the narrower, view of the subject; and historically much of what I am calling macrolinguistics has, therefore, been included within the scope of the term 'applied linguistics'. It requires but little reflection, however, to realize that this point of view is unjustifiable.

We now come to my third point, which follows directly from the conjunction of the other two. This is that, if 'theoretical' is given what one might call its strong, post-Chomskyan, sense, a distinction can be usefully drawn between 'applied' and 'nontheoretical' or, alternatively, between 'first-degree' and 'second-degree' applied linguistics. First-degree applied descriptive linguistics would have as its aim the production of nontheoretical (i.e. unformalized) descriptions of particular languages. Their format would be determined by the practical purposes for which they are intended: they would be purpose-specific, use-related descriptions--descriptions 'for the consumer', if I may echo the title of Michael Halliday's famous *GURT-15* paper (1964). Second-degree applied linguistics, in contrast, would be what is normally referred to as 'applied linguistics': what is applied (to language-teaching, machine-translation, other kinds of automated language-processing, or whatever) would be the products of either theoretical or first-degree applied linguistics. The terms we use to draw these distinctions are relatively unimportant. But the distinctions themselves are highly important. And in drawing them I think I can fairly claim to be doing no more than systematizing, or making explicit, what is implicit in current attitudes towards applied linguistics and its relation to linguistic theory. But in saying this I am of course open to correction by the experts.

But why, it may now be asked, is it not possible to operate with a single notion of the language-system valid in all branches of linguistics, micro- and macro-, theoretical and nontheoretical. The answer derives from the complexity of the phenomena covered by what is pretheoretically identifiable as language (Saussure's *langage*) and the multiplicity of their connections with other phenomena, natural and cultural, constituting the data of other disciplines. There is no immediate possibility, perhaps even no ultimate possibility, of constructing a unified theory of the natural and social sciences within which a unified theory of language (of *langue* construed generically) would find its place and be descriptively and explanatorily adequate to the data that it systematizes.

The ontological status of the language-systems underlying natural languages has been controversial ever since the publication of the *Cours*. Saussure's own views are unclear and perhaps contradictory. At one time, he says that they are supra-individual social facts; at another time, that they are stored in the brains of individual members of the language community (1916: 23-32). Much of the controversy that has surrounded the Saussurean distinction of *langue* and *parole* (or the Chomskyan distinction of 'competence' and 'performance', which is valid for psycholinguistics but not for microlinguistics) turns, in my view, on a false assumption: The assumption that there is only one kind of reality; that natural languages must be either psychological or social entities, or, in terms of an alternative dichotomy, that they must be either physical or nonphysical. What I am saying is that microlinguistics, on the one hand, and the several branches of macrolinguistics, on the other, start from the same pretheoretical notion of what language is and that, according

to their own viewpoint and the alliances that they forge with other disciplines (mathematics, psychology, sociology, anthropology, etc.), they each practice a particular kind of abstraction and idealization in the construction of the ontologically appropriate model of the underlying language-system.

It may be that some day the dream of unified science which inspired, inter alios, Bloomfield, Carnap and Morris, in their time, will again seem realizable. It does not seem so today; and I, for one, as I peer into my glass and see but darkly, would not bet on its being realized this side of eternity. Of one thing I am confident, as I take my view toward the future: it is that premature reductionism of whatever kind, generativist or physicalist, psycholinguistic or sociolinguistic, is detrimental to our subject in its current state of development. And I still think that there is a central place for so-called autonomous microlinguistics, whose objects of study are neither psychological nor social, and which are connected to the primary data, language-utterances, the products of actual and potential language-behavior by methodologically controlled 'rough justice'.

8 Conclusion: Real progress or not? I must now bring my talk to a conclusion by answering the question in my title, and, in so doing, by taking some kind of a 'view toward the future'. Has there been real progress in linguistics over the last 40 years and, if there has been, in what way and to what degree has *GURT* contributed to it? The answers that I, at least, would give to these questions emerge clearly, I think, from what I have said already. But let me make them fully explicit.

Progress implies directionality. It also implies a goal, and perhaps also a series of goals, intermediate and ultimate. But it does not imply that the series is monotonic, objective or intersubjectively constant. Nor does it imply that even what are retrospectively (and Whiggishly) seen as the determinant short-term goals of the participants at any one time are recognized as such, at that time, by the participants themselves. Indeed, paradoxical though it may appear at first sight, it is often easier to be sure of the ultimate goal and to know whether one is significantly closer to it than one's predecessors were than it is to know whether one has made significant progress towards one's declared intermediate, short-term, goals or whether their achievement, if they are achieved, will be retrospectively identified as marking the path of progress.

There can be no doubt that we are closer to our ultimate goal than we, or our predecessors, were 40 years ago. We now have a far better understanding of language and languages, from several legitimately existent and equally valid points of view, than we had then. We no longer confuse the language-system (Saussure's *langue*), from either the microlinguistic or a macrolinguistic point of view, with language-behavior (locutionary, illocutionary, and perlocutionary) or with the transmitted, or transmittable, products of that behavior (Saussure's *parole*). We no longer need to confuse (1) applied linguistics with macrolinguistics or (2) general linguistics with theoretical linguistics. We may still be a long way from having provided a theoretically satisfying answer to the \$64,000 question 'What is language?' from any of the currently recognized distinct, and--I repeat--equally valid, points of view. But we have at least made a start, and not only in microlinguistics (so-called autonomous linguistics), but also in such branches

of macrolinguistics as psycholinguistics and sociolinguistics, with the construction of a theoretical framework within which the answer, or rather the answers, can be formulated. And one consequence of this, of particular concern in the present context, is that we now have a more sophisticated appreciation of the relation between theoretical and non-theoretical micro- and macrolinguistics.

As to nontheoretical first-degree applied descriptive linguistics, progress over the last 40 years has been both self-evident and impressive. Anyone who doubts that this is so has only to look at recently produced commercial dictionaries and grammars, especially of English, but also of other major languages, and to compare them with what was available and counted as authoritative a generation ago. *Si monumentum requiris, circumspice!* It is worth mentioning, also, as a sign of the increased maturity of our discipline that we now know the difference between proper and improper prescriptivism: we are subtler in our presentation of the thesis of linguistic egalitarianism or Panglossia (as Dell Hymes has called it).

The fact that we do undoubtedly have more comprehensive, more accurate and pedagogically more satisfactory descriptions of important world languages than we had 40, or even 20, years ago; the fact that we are better able to relate our descriptions to the varied needs of various groups of 'consumers'; the fact that we can now give a proper, sociolinguistically or psycholinguistically justifiable, defense of neoprescriptivism (i.e. a prescriptivism which, whilst avoiding what I have elsewhere called 'the classical fallacy', recognizes that there are some senses in which not all languages and dialects are equal): all this constitutes progress towards both our ultimate and some of our intermediate goals. And it is in large part attributable to the dialectical relation between theoretical and applied linguistics that has been developed in the major centers of applied linguistics in the United States and abroad (some of them, like the one here in Washington spelt with capitals: CAL), a relation which, as I have emphasized throughout my talk, *GURT* has done much to foster.

When it comes to that part of second-degree applied linguistics upon which *GURT-40*, *GURT 1989*, is focussed--language teaching, testing, and technology--, a taciturn discretion is perhaps the better part of predictive valor. I will leave to other speakers, more competent than I, the task of drawing out the lessons of the past and hazarding 'a view toward the future.'

I will, however, say just a few words about one other area of second-degree applied linguistics, with which I have been concerned in the past: machine-translation. It is now rather more than a quarter of a century since Bar-Hillel dampened the enthusiasm of the first wave of researchers in MT, and of their sponsors, by arguing forcefully and, on his own terms, cogently that what he called Fully Automatic High Quality Translation (FAHQT) was in principle impossible. Research on MT has continued, but at a much reduced level. It has now begun to pick up again and has been integrated with the broader range of applications of linguistics, computer science, and artificial intelligence that is covered by the term 'information technology' and is underpinned, theoretically, by so-called cognitive science.

There is no doubt that progress is being made towards FAHQT, for restricted practical purposes: i.e. for the purposes for which translation, as we

normally understand the term, is usually required. Does this mean, then, that Bar-Hillel was wrong when he made his pessimistic predictions in the 1950s and 1960s? Any answer to that question depends, in part, on one's definition of 'translation' (and 'high quality') and, in part also, on one's philosophical prejudices and preferences. But one thing, I think, is clear: FAHQT, defined as 'fully automatic translation (of certain types of texts) comparable in quality with the best translation produced by competent human translators (which of course is never perfect)' will not be possible until theoretical linguistics, macro- and micro-, is further advanced than it is today and the various branches of linguistics are more satisfactorily integrated with one another to provide a synoptic view of language--synoptic, but, as I emphasized in the preceding section, not reductionist.

And now, as we began with a bit of numerology (light-heartedly, but also with sober intent), so let us end. The next revolutionary change of paradigm in mainstream linguistics, the next major breakthrough, is scheduled for 1996 or 1997: 40 years after the last one. We know that this is so by simple extrapolation, but what that breakthrough will be, we do not as yet know, even though it may well be already in train. What it is will gradually become evident, we can confidently predict by further numerical induction, from about the turn of the century, in the second decade of the next 40 years of *GURT*; and it may formally be declared to have taken place by the 80th president of the LSA, who for all we know is present today. Its effects will gradually work themselves out in the third and fourth decade of the next 40-year period. It may then be that someone will be invited to *GURT-80*, as I have been invited to *GURT-40*, to reflect upon the developments of the preceding 40 years and to put these as yet unforeseen developments in context. If and when that happens, I hope that whoever is charged with that task will be able to look back to my performance, both today and perhaps more generally, and say: 'He didn't do too badly: he got some of it right.'

References

- Aarsleff, Hans. 1970. The history of linguistics and Professor Chomsky. *Lg.* 46.570-85.
 Aarsleff, Hans. 1971. Cartesian linguistics: History or fantasy? *Linguistic Sciences* 17.1-12.
 Aarsleff, Hans. 1982. From Locke to Saussure. London: Athlone.
 Allen, J., B. Patrick, and S. P. Corder, eds. 1973. Readings for applied linguistics. (*The Edinburgh course in applied linguistics*, vol. 1.) London: Oxford University Press.
 Allen, W. Sidney. 1957. On the linguistic study of language: An inaugural lecture. Cambridge: Cambridge University Press. (Reprinted, with Author's note (1965), in: Strevens. 1966:3-26 and (without notes or references), with '[Author's] Foreword [1971]', in: Allen and Corder 1973:147-61.)
 Austerlitz, Robert A., ed. 1975. The scope of American linguistics. The first golden anniversary symposium of the Linguistic Society of America. Lisse, Netherlands: De Ridder.
 Austin, John I. 1962. How to do things with words. Oxford: Clarendon Press and Cambridge, Mass.: Harvard University Press.
 Austin, William M., ed. 1967. Papers in linguistics in honor of Leon Dostert. The Hague and Paris: Mouton.
 Bar-Hillel, Yehoshua. 1953. Some linguistic problems connected with machine translation. *Philosophy of Science* 20.217-25. (Sections 1.3, and 4 reprinted in: Bar-Hillel 1970:306-16; section 2, Inter-translatability of natural languages. Reprinted in: Bar-Hillel 1964:54-58.)
 Bar-Hillel, Yehoshua. 1964. Language and information. Reading, Mass.: Addison-Wesley.

36 / Georgetown University Round Table on Languages and Linguistics 1989

- Bar-Hillel, Yehoshua. 1969. Review of Lyons (1968). In: *Semiotica* 1.449-59. (Reprinted in: Bar-Hillel 1970:364-74.)
- Bar-Hillel, Yehoshua. 1970. *Aspects of language*. Jerusalem: Magnes Press.
- Bar-Hillel, Yehoshua, J. Malino, and A. Margalit. 1964. On logic and theoretical linguistics. In: Sebeok (1974:37-101).
- Bazell, Charles E. 1952. The correspondence fallacy in structural linguistics. *Studies by members of the English Department, Istanbul University* 1952. 138-45. (Reprinted in Hamp et al. 1966: 271-98.)
- Bazell, Charles E. 1964. Three misconceptions of grammaticalness. In: Stuart, ed. *GURT* 1964:3-9. (Reprinted in: O'Brien 1968:25-31.)
- Bazell, Charles E. et al. eds. 1966. In memory of J. R. Firth. London: Longman.
- Benveniste, Émile. 1966. *Problèmes de linguistique générale*. Paris: Gallimard. (English trans. by Mary E. Meek. Coral Gables, Fla.: University of Miami Press.)
- Benveniste, Émile. 1974. *Problèmes de linguistique générale*, vol 2. Paris: Gallimard.
- Bloch, Bernard. 1949. Leonard Bloomfield: Obituary. *Lg.* 25.87-98.
- Bloch, Bernard and G.L. Trager. 1942. Outline of linguistic analysis. (Special publication of the Linguistic Society of America.) Baltimore, Md.: The Waverly Press.
- Bloomfield, Leonard. 1919. Fifty years of comparative philology in America. In: *Transactions and Proceedings of the American Philosophical Society* 50.62-83.
- Bloomfield, Leonard. 1924. Review of Saussure (1916). *Modern Language Journal* 8.317-19.
- Bloomfield, Leonard. 1926. A set of postulates for the science of language. *Lg.* 2.153-164. (Reprinted in: Joos 1957:26-31.)
- Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart and Winston.
- Bright, William, ed. 1965. *Sociolinguistics*. In: *Proceedings of the UCLA Sociolinguistics Conference*, 1964. The Hague: Mouton.
- Bynon, Theodora, and F.R. Palmer, eds. 1986. *Studies in the history of Western linguistics: In honour of R.H. Robins*. Cambridge: Cambridge University Press.
- Carroll, John B. 1953. *The study of language: A survey of linguistics and related disciplines*. Cambridge, Mass.: Harvard University Press.
- Chomsky, Noam. 1955a. Logical syntax and semantics. *Lg.* 31.36-45.
- Chomsky, Noam. 1955b. Semantic considerations in grammar. In: *Georgetown University Round Table on Languages and Linguistics* 1955. Washington, D.C.: Georgetown University Press. 141-53.
- Chomsky, Noam. 1956. Three models for the description of language. In: *IRE Transactions on Information Theory*, IT-2.13-24. (Reprinted, with corrections, in Luce et al. 1965: 105-24.)
- Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
- Chomsky, Noam. 1964a. The logical basis of linguistic theory. In Lunt (1964:914-78). (Revised and expanded, and published separately as Chomsky 1964b; and, in part, in Fodor and Katz 1964:50-118.)
- Chomsky, Noam. 1964b. Current issues in linguistic theory. The Hague: Mouton.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1966a. *Cartesian linguistics*. New York: Harper and Row.
- Chomsky, Noam. 1966b. Topics in the theory of generative grammar. In Sebeok (1966:1-60).
- Chomsky, Noam. 1968. *Language and mind*. New York: Harcourt, Brace and World.
- Chomsky, Noam. 1975a. The logical structure of linguistic theory. New York: Plenum.
- Chomsky, Noam. 1975b. *Reflections on language*. New York: Random House.
- Chomsky, Noam, and G.A. Miller. 1963. Introduction to the formal analysis of natural languages. In: R. Luce et al., *Handbook of mathematical psychology*, vol. 2. New York and London: Wiley, 1963. 269-321.
- Dingwall, W.O. 1966. Recent developments in transformational-generative grammar. *Lingua* 16.292-316.
- Dingwall, W.O., ed. 1971. *A survey of linguistic science*. College Park, Md.: University of Maryland Linguistics Program.
- Dinneen, Francis P. 1967. An introduction to general linguistics. New York: Holt Rinehart and Winston. (Reprinted by Georgetown University Press, Washington, D.C.)
- Dostert, Leon. 1955. The Georgetown-IBM experiment. In: Locke and Booth (1955:124-35).
- Firth, John R. 1957. *Papers in linguistics 1934-1951*. London: Oxford University Press.
- Firth, John R. 1968. *Selected papers 1952-1959* (edited by Frank R. Palmer). London: Oxford University Press.
- Fodor, Jeremy A., and J.J. Katz, eds. 1964. *The structure of language: Readings in the philosophy of language*. Englewood Cliffs, N.J.: Prentice Hall.

- Garvin, Paul L., ed. 1964. A Prague School reader in esthetics, literary structure and style. Washington, D.C.: Georgetown University Press.
- Garvin, Paul L. 1967. The Georgetown-IBM experiment of 1954: An evaluation in retrospect. In: Austin (1967:46-56).
- Gleason, Harold A. 1955. An introduction to descriptive linguistics. New York: Holt, Rinehart and Winston. (Rev. ed., 1961.)
- Gleason, Harold A. 1965. The origins of modern linguistics. In: Linguistics and English grammar. New York: Holt, Rinehart and Winston. 28-47.
- Godel, Robert. 1969. A Geneva School reader in linguistics. Bloomington: Indiana University Press.
- Gray, Louis H. 1939. The foundations of language. New York: Macmillan.
- Hall, Robert A. 1950. Leave your language alone! Ithaca, N.Y.: Linguistica.
- Halliday, Michael A.K. 1964. Syntax and the consumer. In: Georgetown University Round Table on Languages and Linguistics 1964. Washington, D.C.: Georgetown University Press. (Reprinted in O'Brien 1968:11-24.)
- Hamp, Eric, et al. 1966. Readings in linguistics, 2. Chicago: University of Chicago Press.
- Harris, Zellig S. 1951. Methods in structural linguistics. Chicago: University of Chicago Press.
- Harris, Zellig S. 1952a. Discourse analysis. *Lg.* 21.1-30.
- Harris, Zellig S. 1952b. Discourse analysis: A sample test. *Lg.* 28.474-94.
- Harris, Zellig S. 1965. Transformational theory. *Lg.* 41.363-401.
- Hockett, Charles F. 1948. A note on 'structure'. *IJAL* 14.269-71. (Reprinted in Joos 1957: 279-80.)
- Hockett, Charles F. 1954. Two models of grammatical description. *Word* 10.210-31. (Reprinted in Joos 1957:386-99.)
- Hockett, Charles F. 1958. A course in modern linguistics. New York: Macmillan.
- Hockett, Charles F. 1965. Sound change. *Lg.* 41.185-205.
- Hockett, Charles F. 1968. The state of the art. The Hague: Mouton.
- Hoeningwald, Henry M. 1960. Language change and linguistic reconstruction. Chicago: University of Chicago Press.
- Hoeningwald, Henry M. 1974. Fallacies in the history of linguistics and the appraisal of the early nineteenth century. In: Hymes (1974:346-58).
- Hoeningwald, Henry M., ed. 1979. The European background of American linguistics: Papers of the Third Golden Anniversary Symposium of the Linguistics Society of America. Dordrecht: Foris.
- Hoeningwald, Henry M. 1986. Nineteenth century linguistics on itself. In: Bynon and Palmer (1986:172-88).
- Haugen, Einar. 1951. Directions in modern linguistics. *Lg.* 27.211-22. (Reprinted in Joos 1957:357-63.)
- Householder, Fred W. 1952. Review of: Harris (1951). *IJAL* 18.260-68.
- Householder, Fred W. 1957. Rough justice in linguistics. In: Georgetown University Round Table on Languages and Linguistics 1957. Washington, D.C.: Georgetown University Press. 153-73.
- Hymes, Dell, ed. 1974. Studies in the history of linguistics. Bloomington: Indiana University Press.
- Hymes, Dell, and J. Fought. 1975. American structuralism. In: Sebeok (1975:901-1176). (Reprinted separately, with 'Epilogue, 1979'. The Hague: Mouton, 1981.)
- Joos, Martin, ed. 1957. Readings in linguistics. Washington, D.C.: American Council of Learned Societies. (Republished as *Readings in Linguistics* 1. Chicago: Chicago University Press. 1966.)
- Koerner, Konrad. 1989. Practicing linguistic historiography: Selected essays. Amsterdam and Philadelphia: Benjamins.
- Lepschy, Giulio. 1982. A survey of structural linguistics, 2nd ed. London: Deutsch. (1st ed., London: Faber and Faber, 1970, revised from *La linguistica strutturale*, Torino: Einaudi, 1966.)
- Leroy, Maurice. 1971. *Les grands courants de la linguistique moderne*. 2nd ed. Bruxelles: Editions de l'Université de Bruxelles. (1st ed., 1963. English trans. of 1st ed., The main trends in modern linguistics. Oxford: Blackwell.)
- Locke, William N., and A.D. Booth, eds. 1967. Machine translation of languages. New York: Wiley, and London: Chapman and Hall.
- Luce, R. Duncan, R.R. Bush, and E. Galanter. 1963. Handbook of mathematical psychology, vol. 2. New York and London: Wiley.

- Lunt, Horace L., ed. 1964. *Proceedings of the Ninth International Congress of Linguists*. The Hague: Mouton.
- Lyons, John. 1962. Review of: Mohrmann et al. (1961). *American Anthropologist* 64.1118-24.
- Lyons, John. 1968. *Introduction to theoretical linguistics*. London and New York: Cambridge University Press.
- Lyons, John. 1977a. *Semantics*, 2 vols. London and New York: Cambridge University Press.
- Lyons, John. 1977b. Chomsky, 2nd ed. London: Fontana/Collins, and New York: Viking. (1st ed., 1970.)
- Lyons, John. 1981a. *Language and linguistics*. London and New York: Cambridge University Press.
- Lyons, John. 1981b. *Language, meaning and context*. London: Fontana/Collins.
- Lyons, John. 1990. *Essays in linguistic theory*. Cambridge: Cambridge University Press.
- Lyons, John, R. Coates, M. Deuchar, and G. Gazdar, eds. 1987. *New horizons in linguistics 2*. London: Penguin.
- Malmberg, Bertil. 1964. *New trends in linguistics: An orientation*. Stockholm and Lund: The Nature Method Institute.
- Martinet, André. 1953. *Structural linguistics*. In: A.L. Kroeber, ed., *Anthropology today*. Chicago: University of Chicago Press. 574-86.
- Matthews, Peter H. 1979. *Generative grammar and linguistic competence*. London: Allen and Unwin.
- Matthews, Peter H. 1986. *Distributional syntax*. In: Bynon and Palmer (1986:245-78).
- Mohrmann, Christine, F. Norman, and A. Sommerfelt, eds. 1963. *Trends in modern linguistics*. Utrecht and Antwerp: Spectrum.
- Mohrmann, Christine, A. Sommerfelt and J. Whatmough, eds. 1961. *Trends in European and American linguistics*. Utrecht and Antwerp: Spectrum.
- Newmeyer, Frederick J. 1980. *Linguistic theory in America*. New York: Academic Press.
- O'Brien, Richard J., ed. 1968. *Georgetown University Round Table: Selected papers on linguistics 1961-1965*. Washington, D.C.: Georgetown University Press.
- Pedersen, Holger. 1931. *Linguistic science in the nineteenth-century*, trans. J. W. Spargo. Cambridge, Mass.: Harvard University Press. (Republished as *The discovery of language*. Bloomington: Indiana University Press, 1959.)
- Percival, W. Keith. 1976. The applicability of Kuhn's paradigms to the history of linguistics. *Lg* 52.285-94.
- Philological Society. 1978. *The Neogrammarians: Commemorative volume*. Transactions of the Philological Society. Oxford: Blackwell.
- Robins, Robert H. 1964. *General linguistics: An introductory survey*. London: Longman. (3rd ed., 1979.)
- Robins, Robert H. 1979. *A short history of linguistics*, 2nd ed. London: Longman, and Bloomington: Indiana University Press. (1st ed., 1967.)
- Sampson, Geoffrey R. 1980. *Schools of linguistics*. London: Hutchinson, and Stanford, Calif.: Stanford University Press.
- Sapir, Edward. 1921. *Language*. New York: Harcourt, Brace.
- Saussure, Ferdinand de. 1916. *Cours de linguistique générale*. Paris: Payot.
- Sebeok, Thomas A., ed. 1966. *Current trends in linguistics*, vol. 3. The Hague: Mouton.
- Sebeok, Thomas A., ed. 1975. *Current trends in linguistics*, vol. 13. The Hague: Mouton.
- Stuart, C. I. J. M., ed. 1965. *Georgetown University Round Table on Languages and Linguistics 1964*. Washington, D.C.: Georgetown University Press.
- Sturtevant, Edgar H. 1947. *An introduction to linguistic science*. New Haven: Yale University Press.
- Szemerényi, Oswald. 1971. *Richtungen der Modernen Sprachwissenschaft*, Band 1: Von Saussure bis Bloomfield. Heidelberg: Winter.
- Trager, George L. 1949. The field of linguistics. (*Studies in Linguistics, Occasional Papers*, 1.) Norman, Okla.: Battenburg Press.
- Vachek, Josef, ed. 1964. *A Prague School reader in linguistics*. Bloomington: Indiana University Press.
- Vachek, Josef. 1966. *The linguistic school of Prague*. Bloomington: Indiana University Press.
- Waterman, John T. 1970. *Perspectives in linguistics: An account of the background of modern linguistics*, 2nd rev. ed. Chicago: University of Chicago Press. (1st ed., 1963.)

The integration of language and content instruction for language minority and language majority students

G. Richard Tucker and JoAnn Crandall

Center for Applied Linguistics

Our major purpose in this paper is to describe a series of ongoing educational innovations intended to improve the quality of (language) education for language minority and language majority individuals in the United States. We take as our point of departure the observation that the number of school-aged language minority students is increasing dramatically. In many parts of the United States, such students now--or shortly will--constitute a majority of the pupils in local educational agencies (LEAs). In many instances, students enter school with minimal or no proficiency whatsoever in English. In other instances, language minority students who seem--at least to all outward appearances--proficient in social language skills often have difficulty in acquiring the academic language and the cognitive skills which they need for success in their math, science, and other academic subjects. The purpose of our paper is not to reenter the debate on the optimal (language) educational strategies for such youngsters; others have done that (e.g., Hakuta 1986; Willig 1985). Rather, we wish to describe an emerging educational practice that seems to offer great promise for such students.

Simultaneously, large numbers of language majority children participate in sequences of foreign language study at the elementary or secondary level without apparently ever developing meaningful proficiency in their target language. A nationwide survey of foreign language enrollments by CAL staff (Rhodes and Oxford 1988) revealed that approximately 22% of our nation's elementary schools and approximately 87% of our secondary schools offer programs of foreign language instruction. However, the best guess we can make is that fewer than 1% of the students who are enrolled in such programs--already a relatively small number of youngsters--participate in programs in which the development of bilingual proficiency is either a demonstrable program goal or an attainable objective. (In passing, it should also be noted that our students, for the most part, do not study any of the so-called less commonly taught languages--languages which are after all spoken by a large majority of the world's population.) Is it realistic for language majority children to acquire bilingual proficiency by participating in foreign language programs within our public school system? Again, the purpose of our paper is not to examine the broad array of methods and approaches used to deliver foreign language instruction (cf. Larsen-Freeman 1986; Richards

and Rodgers 1986; Schinke-Llano 1985; Stern 1983). Instead, we wish to describe an educational practice that seems to hold great promise for improving the quality of foreign language instruction, and concomitantly the degree of proficiency attained.

The integration of language and content instruction. During the last five years, increasing attention has focused on what has come to be called 'content-based language instruction' (sometimes referred to as language-sensitive content instruction when implemented by the content-area teacher). Such instruction offers the possibility of broadening and deepening the language proficiency of students of foreign and second languages. A number of encouraging models of integrated language and content instruction have been tried at elementary, secondary, and tertiary levels. These include sheltered English classes, classes in specific academic language skills, foreign language immersion programs which teach content in the second language, and parallel courses in language and content (see for example, Cantoni-Harvey 1987; Chamot and O'Malley 1987; Crandall 1987; Mohan 1986; Short, Crandall and Christian 1989; Snow and Brinton 1988; Willetts 1986). Recently, Snow, Met, and Genesee (1989) have presented a conceptual framework for integrating language and content instruction which they illustrate with examples from both foreign language and ESL classes.

An approach which seeks to integrate language and content instruction contrasts sharply with many existing practices or methods in which language skills are taught virtually in isolation from substantive content. Often, the use of traditional approaches means that the target language is not used for communication in any meaningful, purposeful, or even natural contexts. It is the feeling of those associated with this move toward integrating language and content instruction that the substantive content of a curriculum provides both a motivational and a cognitive basis for driving or for enhancing language learning--and that (target) language development and cognitive development can best proceed simultaneously. Another rationale for the strongest possible integration of language and content instruction lies with the necessity for the child to acquire the "school register" of language--that is, to acquire proficiency in the cognitive academic language which is needed to participate effectively in cognitively demanding, complex, and often abstract school learning. Although this need has been cited as a matter of particular concern for teachers of limited English proficient (LEP) youngsters, it is certainly also a concern for those who seek to develop higher level skills in foreign languages for language majority youngsters.

A brief diversion. Although we did not realize it at the time (and certainly did not label it as such), we began exploring the area of content-based language instruction at McGill University in the mid-1960s. At the time, a number of us--parents, educators, and researchers (see, for example, the Introduction to Lambert and Tucker 1972)--were becoming increasingly concerned by the fact that anglophone or English-speaking youngsters could somehow participate in 12 years of foreign or second language instruction in French; pass the required matriculation examinations set by the Ministry of Education of the Province of Quebec; but still not function fluently and

confidently in educational, occupational, or social settings following their graduation. This same story was repeated--and perhaps with even less tangible results--in the other nine Canadian provinces and two territories. In the United States, of course, virtually no public school graduates whatsoever developed target language proficiency by following any regular sequence of courses offered by LEAs throughout the country.

At McGill University, several of us proposed exploring the implementation of innovative educational programs that might lead to the development of full and effective bilingual proficiency on the part of public school participants--programs that might lead to the development of social and academic language skills in both the first and the second language, and that might result in the acquisition of the requisite content material that students were expected to cover. Thus, in 1965 we began a pilot program in the community of St. Lambert, Quebec in which anglophone children were introduced to their early schooling principally through the medium of French--a second language for them. The so-called Canadian immersion studies have been described in great detail in many places (cf. Genesee 1987; Lambert and Tucker 1972; Swain 1984). Briefly, the salient characteristics of these programs have been that:

1. The programs were designed for English-speaking youngsters in response to continuing parental dissatisfaction with the level of French attained by children participating in English instructional programs with French-as-a-second language components.
2. Participation in such programs has always been voluntary.
3. Parents played an extremely strong and catalytic role in all aspects of program design, development, and implementation.
4. French was used from the very beginning for all readiness activities and as the language of initial reading instruction. For the most part, the native language and target language skills of entering children were uniform.
5. The teachers were native speakers of French, which was used as the major medium of primary content instruction. Nevertheless, an English Language Arts component was always added to the curriculum by grade 2 or 3 and formed an integral part of the core curriculum throughout the primary and secondary years of instruction. This intentional teaching of the mother tongue served to mark explicitly the continuing importance or status of English as a valued language and helped to solidify formal decontextualized, or academic, English language skills.
6. Starting with the middle primary years, some content material was taught via English. Selected periods of the day were devoted from that point on to the teaching of English Language Arts, French Language Arts, some content subjects via French and other content subjects via English.

Such programs are *not* transitional bilingual education programs. Neither are they the traditional second or foreign language programs in the American sense. Rather they are characterized by the use of the target language as a

major medium for teaching relevant content material which is of interest to the students; they contain structured language arts components in both the target language and in the child's mother tongue; they involve continuing communication between the language arts teacher and the content instructors; they are characterized by teaching strategies which are designed to facilitate the provision of comprehensible input to the children; they are designed to maximize the use by teachers and students of diverse strategies for the negotiation of meaning; and they require that students produce varied oral and written output.

The results of research conducted with immersion programs have been presented in some detail, and it would be redundant to repeat them. However, by way of brief summary, the consensus of researchers who have worked with large groups of different children participating over long periods of time in programs in different schools, within different school boards, and in different provinces--and coincidentally in the United States immersion programs as well--is that this innovative approach to second language teaching in which the target language is used as a major medium of classroom communication facilitates second language acquisition without causing any detrimental effects whatsoever to native language development, or to general cognitive or social development. In addition, youngsters perform as well as their English-taught peers on achievement tests in content areas such as math, science, or social studies while performing as well as French-taught peers on comparable French-language achievement tests. In addition, there appears to be evidence for a general increase in creativity or cognitive flexibility associated with the development of bilingualism and with participation in such programs (results which are not inconsistent with those reported by Hakuta 1986).

Thus, there has been a systematic tradition of integrating language and content instruction within the Canadian educational context for more than 20 years. Immersion programs were "exported" to the United States following Russ Campbell's brief sabbatical visit to McGill University during early 1970. Following this visit, Campbell and his colleagues implemented the Culver City Spanish immersion program and the model was subsequently replicated in numerous sites throughout the United States (see Campbell et al. 1985; Tucker and Crandall 1985). The general pattern of findings from immersion programs in the United States has been consistent with the Canadian experience. What critical component was missing from the Canadian program?

A 20-year paradox: The absent peer group. On numerous occasions, when Canadian-based staff would talk about the French-immersion programs to American audiences, a question often raised was why francophone youngsters had not been included so that the classes would truly become two-way bilingual programs. There is no simple answer to this question although there were a variety of religious and political factors which precluded developing what might be referred to as two-way or interlocking or bilingual (immersion) education programs. Suffice it to say, that except for a brief "experiment" in which anglophone youngsters participating in the French-immersion programs were literally transported in groups by taxi across town to participate in selected classes during a part of the school day in a traditional

French Protestant school it was unfortunately not possible to implement programs of this type. Nevertheless, on the basis of this brief experience in the early 1970s, and on the basis of our own continuing intuitions, many of us long felt strongly that truly innovative programs which brought together mother tongue speakers of the two major contact languages for meaningful instruction in each of those languages would ultimately provide the greatest opportunity for children to develop full and effective bilingual proficiency.

Thus, a number of us had been flirting with a special kind of integration of language and content instruction for more than two decades. We had stopped short, however, of establishing two-way or bilingual immersion programs in which language and content instruction was delivered in a full bilingual program bringing together representatives of both target language groups. But we had written about the potential value of "enrichment models" (Tucker 1986) or programs which could be designed to capitalize on the fact that language-minority students and language-majority students can participate meaningfully and effectively in joint education. Careful, empirical, longitudinal research has demonstrated repeatedly that the facilitation of bilingualism is associated with intellectual advantage, and greater awareness and tolerance for ethnic diversity. An approach which maximizes the integration of language and content instruction for members of major language contact groups simultaneously seemed to hold great promise for building and sustaining valuable natural language resources within the United States (see Campbell and Lindholm 1987; Lindholm 1987).

Two-way or bilingual immersion programs. We believe that there is an emerging awareness of the power and the possibilities of bilingual immersion programs. Such programs have been referred to as two-way, interlocking, dual-language, bilingual immersion, or by some as developmental bilingual programs. Our operational illustration of an exemplary bilingual immersion program follows. Let us suppose, for example, that there are 30 youngsters in a particular grade 1 class at a typical elementary school. For illustrative purposes, let us assume that 15 of them are Anglo or English mother tongue youngsters, and 15 are Hispanic youngsters (the language background characteristics of the participating Hispanic youngsters will, in many cases, be more heterogeneous than that of their Anglo counterparts). The youngsters would be together in a class in which some portion of the day would be devoted to Spanish language arts (for the Hispanics), Spanish as a second language (for the Anglos), English language arts (for the Anglos), English as a second language (for the Hispanics), with the teaching of selected content material--let us say mathematics--in English, and other content material--let us say history--in Spanish. Over the course of several years, the idea would be to offer a program of bilingual instruction in which representatives of both of the ethnolinguistic groups would have an opportunity to develop and to hone their literacy skills while developing the fullest possible social and academic proficiency in both of their languages.

In addition to the regular continuing sequence of language arts classes, care would be taken so that children would have an opportunity to study all of the content subjects in both of the languages during the course of their school experience. This would facilitate the development of the appropriate

"registers" for mathematics, science, etc. in each of the two languages. Instruction would be offered within a bilingual ambience in which teachers as well as students would be available to provide good language models and maximize the opportunity for cooperative learning, peer group tutoring, etc. Do such exemplary programs exist?

Exemplary bilingual immersion programs. Lindholm (1987) compiled a list of the preschool through high school bilingual immersion programs in the United States. At that time, she identified 60 such programs in California, Florida, Illinois, Massachusetts, Michigan, New York, Virginia, and Washington, D.C. The most prevalent languages involved were English and Spanish, although there were also programs involving English and Greek (in Long Island) and English and Arabic (in Hamtramck, Mich.). Lindholm discusses the rationale for bilingual immersion education and presents a theoretically motivated operational definition of such programs. She describes each program in some detail and provides directions for obtaining additional information. According to Lindholm, the following criteria are essential for successful bilingual immersion programs:

1. The bilingual instructional treatment must be provided for at least 4 to 6 years (note how different this is from the typical 'early exit' transitional bilingual education program).
2. There will be a focus on the regular academic curriculum as well as on language development.
3. There will be the fullest possible integration of language arts with the curriculum.
4. Optimal dual language input should be provided through communicatively sensitive language instruction and subject-matter presentation.
5. There will be ample opportunity and demand for language output.
6. Instruction will be carried out in what might be referred to as an "additive" bilingual environment.

Thus, she identifies a set of criteria which are fully compatible with those described by Snow, Met and Genesee (1989), Short, Crandall and Christian (1989) as well as with the earlier theoretical construct of Mohan (1986).

Two case studies. Let us describe briefly two exemplary bilingual immersion programs. We have chosen these two illustrative programs because they were implemented under quite different social and ethnolinguistic circumstances in two widely separated parts of the country and, equally importantly, because each has been the subject of careful attention since inception. It is our belief that additional longitudinal quantitative and qualitative evaluation is necessary and desirable to document the relative efficacy of bilingual immersion and to describe the conditions under which it is a viable program option for youngsters.

Case Study 1: Key Elementary School (Arlington, Va.). Under the aegis of the federally funded Center for Language Education and Research (CLEAR), CAL staff have worked with staff of the Key Elementary School

in Arlington County, Va., to design and to implement a bilingual immersion program. The program is currently in its third year of operation. The pilot class of youngsters is now in the third grade, with new cohorts having been added in each of the two succeeding years. (Children participated in a "regular" English-medium kindergarten program.) The County curriculum has been used for all content areas so that students in this program work toward the same objectives as other elementary school youngsters in Arlington County schools. That is, the students follow the typical program in social studies, science, mathematics, and language arts. Participation in the program is voluntary. The classes average from 18 to 20 students with usually about 40% being native Spanish speakers, 40% being native English speakers, and 20% being speakers of other languages who have nevertheless added English to their repertoire. The Spanish-speaking youngsters come from disparately, but predominantly, Central American backgrounds.

The program has been a bilingual one using English and Spanish as the media of instruction. Teachers who are native speakers of Spanish provide instruction during the "Spanish portion" of the day while the converse is true for the English portion. The objective was to devote about 50% of the available instructional time to Spanish, and 50% to English. An important contributing factor in the successful implementation of the program has been the support and encouragement of the school principal Dr. Paul Wireman, and enthusiastic support by parents.

CAL staff (in particular Nancy Rhodes, with assistance from Donna Christian and Jodi Crandall) have followed the program closely since its inception. In general, their evaluation plan addressed the following general questions:

What level of English and Spanish proficiency do students attain, and how does this change over the year?

How well do the bilingual immersion students perform in content area subjects? Do they appear to make academic progress comparable to other students following the regular curriculum? (There is a potential comparison group, but not a matched control group in the strict sense of the term.)

How do affective factors (cross-cultural awareness, language attitude, etc.) change over the course of participation in the program?

How do parents view the program?

Information of various types has been collected over the past three years (data for year 3 are, of course, not yet complete, but will be collected later in the spring). In addition to collecting quantitative data, staff members have also spent time observing classes and talking informally with various staff. Students, teachers, parents, and the principal were also interviewed to elicit their feelings about the program. The results can be briefly summarized as follows (more detailed information is presented in Center for Applied Linguistics 1988).

(1) With respect to English language development (as assessed by the Language Assessment Scales (LAS), a Student Oral Proficiency Rating (SOPR), and the Boehm Test of Basic Concepts) the English-speaking

students do improve both their Spanish and English skills from fall to spring of grade 1. By grade 2 the English speakers had reached the top of the scale in their English, and had improved their Spanish quite noticeably. The Spanish-speaking youngsters also improved from fall to spring and by the end of the 2nd grade performed at nearly the highest levels in both Spanish and English. On the basis of the Boehm scores, the immersion class students showed gains in their understanding of concepts from fall to spring in both English and Spanish--and additionally in both year 1 and year 2 of the evaluation the immersion class students had gains equal to or higher than a comparable paired first grade class. (This was, as mentioned before, not a matched control group in the traditional sense nor were students randomly assigned to classes.) In terms of English reading progress, by the end of second grade, 16 of 18 immersion students scored at the first half of the third grade level, according to tests that accompany the Ginn reading series.

(2) With respect to social studies which was taught in Spanish and tested in Spanish, there was virtually no difference in scores when compared with class averages for the paired classes that were taught and tested in social studies in English. Likewise, students appeared to be making appropriate gains in Spanish and in mathematics. For math, the students were administered a test in English developed for the Holt math series, with the majority scoring between the 80th and 95th percentile.

(3) The students' personal views about learning Spanish and English are also interesting. We were amazed at how sophisticated their knowledge was about language learning and about the differences between the languages. All students enjoyed being able to communicate with others in Spanish and in particular, by grade 2, English students were quite relaxed about their Spanish language skills and seemed to take it for granted that they could speak two languages. The anglophone children suggested that "Of course Spanish was useful--that's how we can talk to people who only speak Spanish."

(4) In general, parents uniformly spoke extremely favorably about the program. In particular, parents with several children in the County's schools have noted that academically the participants seemed advanced compared to other children. Parents believe that the program is excellent both academically and socially.

All in all, this particular program appears to be a success. The children appear to interact; to develop social and academic language skills in both English and Spanish; to gain content mastery in those subjects which are a part of the regular County curriculum; and in short, to have benefited from this program in which there is an optimal integration of language and content instruction. Success of the program is, at least in part, attributable to the excellent coordination among staff and to the support of the administration. CAL staff will continue to monitor the program at the request of school officials.

Case study 2: Edison Elementary School (California). The second case study is taken from work conducted by CLEAR staff, in particular by Kathryn Lindholm, with the Edison Elementary School in the Santa Monica-Malibu Unified School District in California. A complete report is presented by Lindholm (1988) in which she summarizes the language proficiency and

academic achievement of two cohorts of students after their first year of participation in a bilingual immersion program.

Briefly, in this particular school, English-speaking and Spanish-speaking youngsters--58 at the kindergarten level and 54 at the first grade level--participated in a bilingual immersion program which was slightly different from that implemented at Key Elementary School. According to the instructional design for this program, both native English-speaking and native Spanish-speaking kindergarten and first grade students received approximately 90% of their instructional day in Spanish. During that portion of the day Spanish Language Arts was taught and in addition all content material was taught in Spanish. For the remaining 10% of the day, teaching was carried out in English by another teacher. This time was used for English Language Arts and physical education. The instructional content of the program was equivalent to that for other nonparticipating students in the school. Enrollment in the bilingual immersion program was voluntary (and students were not randomly assigned to classes for experimental purposes). The program integrated language instruction with content instruction in the traditional academic areas and provided for "integrated" classrooms with respect to students' relative language proficiency. Parents were involved in a positive collaborative relationship with teachers and administrators, and efforts were made to develop a bilingual social ambience for the program within the school.

CLEAR staff, under the direction of Lindholm, conducted research intended to examine various aspects of student language characteristics, language proficiency development, achievement in content subjects in both Spanish and English, perceived self-competence, and attitudes toward the bilingual immersion program. The results can be divided into several clusters. With regard to students' language proficiency, *all* made gains in *both* languages. (The Comprehensive Test of Basic Skills, the Idea Proficiency Test, and the Student Oral Language Observation Matrix developed by the California Department of Education were employed.) Native language proficiency was high for members of both groups. The level of proficiency in the second language varied considerably, ranging across the spectrum from nonproficient to fully proficient. Interestingly, more Spanish-dominant students were fluent in the second language than were English-dominant students.

Both groups of students scored at an average to an above-average level in terms of their achievement on tests of reading ability and mathematics ability. The children made significant progress from the fall to the spring. On English language achievement tests (and remember that the students had studied in Spanish) the first grade Spanish-dominant children scored only slightly below average, and they did not differ significantly from the non-bilingual immersion students. The first graders scored well above average, on the other hand. (Measures such as La Prueba Riverside de Realización en Español, CTBS, and CTBS-Español were used.) All of the students' perceived competency ratings were high in each of the domains--cognitive, physical, peer, and maternal--and attitudes toward the bilingual immersion program were generally positive from the parents and from the teachers.

In general, the Spanish speakers at both grade levels made highly significant gains in English and the English speakers demonstrated gain in English proficiency as well. Thus, despite the small amount of English instruction, most student nevertheless made gains in English language proficiency. Second, all students made gains in Spanish proficiency. These are important results because they help to reinforce the importance of the bilingual immersion model's assumptions relating to language development. Interestingly enough, the English-speaking youngsters acquired enough content after only one year of instruction through Spanish to be able to score average to above average in a test normed for native Spanish speakers. The fact that all students were able to score this high in reading and math demonstrates that the students were acquiring the math concepts in Spanish and were able to transfer and apply them when tested in English. Thus, at least tentatively, the achievement results validate the assumptions underlying the bilingual immersion model, which assumes transfer of skills across languages. At the end of grade 1, and now continuing of course into grade 2, the program appears to be an effective education model for both language minority and language majority students. In this particular program the plan is to gradually increase the amount of English content instruction until the program is approximately evenly balanced in this regards. What generalizations, if any, can be drawn on the basis of these two case studies?

The case studies in perspective. Although we have not done justice to the richness of the two case studies in this brief presentation, we hope that they do serve to illustrate the following points. Bilingual immersion education can be a powerful vehicle for promoting the development of bilingual language competence in elementary school-aged youngsters (and presumably for older students as well). This innovation, which represents a "special case" of the integration of language and content instruction, serves to foster the development of solid building blocks in both languages which can lead to the development of social as well as academic language skills (or, as others would label these terms, to the development of contextualized and decontextualized language abilities). The success of the innovation appears to rest upon a sensitivity by the teachers to the language abilities and needs of the children; to the fact that content which is inherently interesting and appealing for the children can be a conceptual peg upon which to build the development of language and higher order thinking skills; and that students working collaboratively across language boundaries can serve to reinforce, to extend, and to solidify their respective language skills. Thus, from our perspective, sensitive teachers working within a positive ambience--in this particular case with supportive parents as well as supportive administrators--can utilize the natural resources which both groups of students bring with them to the learning environment. These abilities can be nurtured and can be extended by careful planning and by creative and sensitive teaching; but the children themselves play a key role in fostering and facilitating this cross-language development.

In a companion paper (Crandall and Tucker 1989), we describe some of the critical attributes which are essential for the successful implementation of programs to integrate language and content instruction. We also identify

areas for further development, for in-service education, and areas in which additional research needs to be conducted (for example, what is the optimal mix of youngsters by language background in a bilingual immersion program; what proportion of instruction should be offered in each language at various levels; cumulatively how do participating children fare when compared with English-instructed peers).

On previous occasions, several of us (see, for example, Campbell and Lindholm 1986; Tucker 1986) have referred to language competence as a national natural resource to be nurtured and sustained. Bilingual immersion education, which is a special case representing the fullest extent possible of integration of language and content instruction, would seem to offer a powerful vehicle for accomplishing this goal.

References

- Campbell, R.N., T.C. Gray, N.C. Rhodes, and M.A. Snow. 1985. Foreign language learning in the elementary schools: A comparison of three language programs. *The Modern Language Journal* 69. 44-54.
- Campbell, R.N., and K.J. Lindholm. 1987. Conservation of language resources. (Educational Report No. 6) Los Angeles: University of California, Center for Language Education and Research.
- Cantoni-Harvey, G. 1987. Content-area language instruction: Approaches and strategies. Reading, Mass.: Addison-Wesley.
- Center for Applied Linguistics. 1988. Review of the second year of the partial immersion program at Key Elementary School, Arlington, Va. (1987-88). Center for Language Education and Research.
- Chamot, A.U., and J.M. O'Malley. 1987. The cognitive academic learning approach. *TESOL Quarterly*, 21(2), 227-47.
- Crandall, J., ed. 1987. ESL through content-area instruction. Englewood Cliffs, N.J.: Prentice-Hall Regents/Center for Applied Linguistics.
- Crandall, J., and G.R. Tucker. 1989. Content-based language instruction in second and foreign languages. Paper presented at RELC Regional Seminar, Singapore.
- Genesee, F. 1987. Learning through two languages: Studies of immersion and bilingual education. Rowley, Mass.: Newbury House.
- Hakuta, K. 1984. Mirror of language: The debate on bilingualism. New York: Basic Books.
- Hamayan, E. 1986. The need for foreign language competence in the United States. ERIC Digest. Washington, D.C.: Center for Applied Linguistics.
- Lambert, W.E., and G.R. Tucker. 1972. The bilingual education of children. Rowley, Mass.: Newbury House.
- Larsen-Freeman, D. 1986. Techniques and principles in language teaching. New York: Oxford University Press.
- Lindholm, K. J. 1987. Directory of bilingual immersion programs: Two-way bilingual education for language minority and majority students. (Educational Report No. 8) Los Angeles: University of California, Center for Language Education and Research.
- Lindholm, K.J. 1988. The Edison Elementary School bilingual immersion program: Student progress after one year of implementation. (Technical Report No. 9) Los Angeles: University of California, Center for Language Education and Research.
- Mohan, B.A. 1986. Language and Content. Reading, Mass.: Addison-Wesley.
- Rhodes, N.C., and R. Oxford. 1988. Foreign language in elementary and secondary schools: Results of a national survey. *Foreign Language Annals* 21. 51-69.
- Richards, J.C., and T.S. Rodgers. 1986. Approaches and methods in language teaching: A description and analysis. New York: Cambridge University Press.
- Schinke-Llano, L. 1985. Foreign language in the elementary school: State of the art. Orlando, Fla.: Harcourt Brace Jovanovich.
- Short, D.J., J. Crandall, and D. Christian. 1989. How to integrate language and content instruction: A training manual. (Educational Report No. ____) Los Angeles: University of California, Center for Language Education and Research.

50 / Georgetown University Round Table on Languages and Linguistics 1989

- Snow, M.A., and D.M. Brinton. 1988. Content-based language instruction: Investigating the effectiveness of the adjunct model. *TESOL Quarterly* 22. 553-74.
- Snow, M.A., M. Met, and F. Genesee. 1989. A conceptual framework for the integration of language and content in second/foreign language instruction. *TESOL Quarterly*.
- Stern, H.H. 1983. *Fundamental concepts of language teaching*. Oxford: Oxford University Press.
- Swain, M. 1984. A review of immersion education in Canada: Research and evaluation studies. In: P. Allen and M. Swain, eds., *Language issues and education policies*. Oxford: Pergamon.
- Tucker, G.R. 1986. Implications of Canadian research for promoting a language competent American society. In: J.A. Fishman, ed., *Festschrift for Charles A. Ferguson*. The Hague: Mouton.
- Tucker, G.R. 1986. Developing a language-competent American society. In: D. Tannen, ed., *Georgetown University Round Table on Languages and Linguistics 1985*. Washington, DC: Georgetown University Press.
- Tucker, G.R. and J. Crandall. 1985. Innovative foreign language teaching in elementary schools. In: P.H. Nelde, ed., *Methods in contact linguistic research*. Bonn: Dümmler.
- Willetts, K. 1986. Integrating language and content instruction. (*Educational Report No. 5*) Los Angeles: University of California, Center for Language Education and Research.
- Willig, A.C. 1985. A meta-analysis of selected studies on the effectiveness of bilingual education. *Review of Educational Research* 55. 269-317.

Practice makes less imperfect: Users' needs and their influence on machine translation development

Veronica Lawson
London, England

Abstract. The problems which attract machine translation developers are not necessarily those that loom large in practice. Only by exercising a system on large quantities of normal text, as at Georgetown in the 1950s, can it be made to reflect the needs of customers.

After a brief historical summary, the relationship between practice and development is explored with reference to 'teaching', testing and technology. Ways are described in which systems were improved to take account of users' needs. The evaluation of machine translation is discussed, as is technological progress in the field.

Practice is the best of all instructors.
Publilius Syrus, ca. 42 B.C.
And practice drives me mad.
Elizabethan MS., 1570

These two quotations in a sense sum up this paper on the relationship between practice and development in machine translation. Specifically, it deals with the needs of customers, and how those needs are reflected in systems.

For this Georgetown University Round Table it is a particularly appropriate subject, for two reasons. Firstly, the Round Table is celebrating thirty-five years of machine translation. As you know, the first ever demonstration of machine translation on an electronic computer was the Georgetown/IBM experiment in 1954. The approach to machine translation here was pragmatic, firmly rooted in real text translated for real users. Georgetown dealt with the truly natural language found in 'translation as she is paid for', rather than the examples that occur to researchers. Not many people will pay for translations of *Time flies like an arrow* or even *The cat sat on the mat*.

Secondly, this Round Table is concerned with 'Teaching, Testing, and Technology'. All of these have been crucial in machine translation (MT): teaching because a computer must be 'taught' about language; testing because this is a major factor in MT's development and use; and technology because it not only permitted MT in the first place, but has shaped the course it has followed ever since.

History. Machine translation (MT) is translation generated by a computer, with or without human assistance (usually with it). In the mid

1950s, when machine translation began, the computer was still an intriguing new tool. It had mastered numbers with intoxicating speed, and researchers expected it to master words with something of the same alacrity. They were of course wrong, and in 1965 their innocent optimism gave way to disillusion when a committee set up by the U.S. National Academy of Sciences advised against further research into machine translation. This was the Automatic Language Processing Advisory Committee, whose famous report (ALPAC 1966) drastically cut government funding for MT research, not only in the United States but to a considerable extent elsewhere in the world. The study now appears questionable in a number of respects, as we shall see later.

Machine translation's second decade, therefore, was quiet. A few MT teams survived in the United States, the Soviet Union, France, and Germany, but they were usually small and on a low budget. The world at large, however, was changing. There was the continuing 'information explosion' (the great expansion of scientific and technical information which had begun in the 1950s). There were also large-scale increases in trade and international cooperation, and a rise in linguistic nationalism in bi- and multilingual regions. All combined to produce a 'translation explosion'. Some government institutions have used MT extensively since before 1970.

From 1975, therefore, in the third decade of MT, there was a renaissance of machine translation, first in Canada and Europe, then in Japan. Quality has improved considerably, and now even the United States government is investing in MT again. Whereas thirty-five years ago the pressure for MT had come from researchers with an intoxicating new tool, the driving force since the mid 1970s has been the users' need for faster and ever more translation. (For a scholarly and readable history, see Hutchins 1986.)

Users' needs affect MT systems in small ways as well as large. This paper sets out a few of the changes which have resulted. Some examples come from my own development work, others from users' experience of various practical MT systems. Since the changes affect teaching, testing, and technology, it may be useful to look at them under those headings.

Teaching. First, teaching: we must 'teach' language to the machine, as far as we can when our own knowledge is but sketchy, and a truly MT-oriented linguistics still eludes us. Linguistic software is powerful, and MT dictionaries include far more kinds of information than dictionaries for human use. A good machine translation system is therefore an Expert System, incorporating the linguistic insights of both its developers and the practitioners who exercise it.

If a machine translation system is to cope with the demands of practice, it is essential to do as Georgetown did in the 1950s: process large quantities of real text, and address the problems thrown up. The more the system aspires to be what my typology classifies as a 'try anything' system (Lawson 1982), the greater the quantity and variety of text must be.

In such a system the rules 'taught' to the software and dictionaries must be as general as possible, to be applicable to many sorts of text. My first MT study was a feasibility study on the machine-translatability of patents, performed in 1979/80 for the Commission of the European Communities on their English-French and French-English Systran systems, then young. One

of my most surprising discoveries in that study was that there was no combined index of all the Systran dictionaries. I actually had to recommend that one be created for the development team. The Systran dictionaries were both a strength and a weakness: a strength, because they were powerful and flexible enough to deal with most of the difficulties which arise in natural language; but a weakness, because they were insufficiently classified and tended to grow excessively and in too random a manner. We went to great lengths to contribute dictionary entries as 'universal' as possible, but the lack of a combined index made it difficult and sometimes impossible to find groups of related expressions and substitute a more general rule, and to discover some sources of error.

The problems which loom large in practice are often not those anticipated by the developers. My patent study got off to an inauspicious start when the translations would not run at all. After six months it was found that the handling program regarded 'unusual' characters and improbably long sentences as errors, and that it rejected entirely any corpus containing ten or more errors. Sentences of more than 105 words were regarded as improbable. Unfortunately, sentences in patents and other legal documents frequently run to half a page, sometimes to a page or even more. Similarly, 'unusual' characters included percent signs and mathematical symbols, very common in patents and other technical and even commercial documents. The maximum sentence length was therefore increased to 255, and 'unusual' characters were made acceptable.

Some of the more obvious problems--problems that one might feel any MT system must be able to solve--may be best left to the posteditor. Questions, for example, may be processed very badly in an MT system, and so researchers may spend much effort on them. However, they are surprisingly uncommon in many translated documents, and entirely absent from patents and some other text types. There is a limit too to the effort worth expending on the exceptionally complex area of the article, about which long books are written, and (short) wars fought: the excuse for the Six Day War between Israel and Egypt was the discrepancy between *territories* and *les territoires* in the parallel texts of a treaty. It is safer and more economic for the posteditor to check and amend articles than for the developer to deal with them exhaustively.

Numerals, on the other hand, are often troublesome, especially where mixed with text. Dates, for example, tend not to be addressed adequately, if at all, by developers before their system is implemented in a translation service. They then discover that dates are found frequently and in notable variety: *March 10* or *10th, 1989*; *10* or *10th March 1989*; *3.10.89* or *03.10.89*, or the European variants with transposed day and month; and so on. The Logos MT system will translate American-style *3.10.89* to European *10.3.89*; this can backfire, however, when a 'date' is detected in error.

Once early in the life of the European Communities' English-French Systran system, *U.S. Patent 1 234 567* became *1 Brevet des Etats-Unis 234 567*. The translators, and hence the system, had had instructions to expand the premodifier *U.S.*; and the patent number in the source text contained spaces instead of commas to mark out groups of three digits. Usually, punctuation must be preserved, but sometimes, to convey the same information, it must be

translated into the punctuation appropriate to the target language. Here French uses spaces; and at that time the commas in English numbers had proved so problematic that they were replaced by spaces when the source text was keyed in, as it then was--a rare and probably short-lived example of the MT tail wagging the language dog. At the time the problem was solved by translating *U.S. Patent* as *Brevet U.S.*

Like a space or a comma, a period, of course, can fulfill various functions. The software may need to interpret it differently according to whether, instead of marking the end of a sentence, it signals a decimal point (in which case it may need translation to a comma), an abbreviation, ellipsis (if one of three dots), another mathematical symbol, etc. If the software breaks up a sentence by mistake at an abbreviation, for example, it will produce nonsense. At least one fledgling system even used periods to mark the ends of lines. This could make for difficulties when the line did not contain a sentence and the software therefore could not find a verb form. My favorite MT error was in the address on a letter (Morgan-Girard, p.c. 1981):

ORIGINAL	John Smith. Managing Director.
RAW MT	John Smith. Directeur se débrouillant.

A director managing, or coping. Normally, the software translated the expression *managing director* correctly. On that occasion it missed it, looked for a verb, and found a job description.

Such howlers are fun, but it should be stressed that they are not the normal run of MT. Much more typical is the following, from a random European Community document which I fed to Systran:

ORIGINAL	The problems we are to consider are difficult ones and will not be easily resolved.
RAW MT	Les problèmes que nous devons considérer sont des difficiles et ne seront pas facilement résolus.

A major problem faced by MT developers when they 'go public' is the preservation of format, which is often a vital component in the transfer of meaning. My first MT study ten years ago was criticized for regarding format and punctuation as part of language, but within a couple of years the critic was saying the same. Format is, I believe, a visual representation of the underlying structure of a text. It helps to make logical connections explicit. Thus we tend to slow down when faced with a text which is presented sentence by sentence instead in paragraphs (e.g. when source and target texts are displayed side by side); and many of us would prefer wordprocessor screens to be full page size. To tamper with format is unwise. Not only may it delay the reader's understanding. Worse, it may interfere with analysis, particularly when text is arranged in columns. Finally, it may seriously inconvenience the customer if, as in the computer manuals which are classic subjects for machine translation, the text is to be printed with illustrations.

Another area of some importance is the postediting of inflections. It may be the work of a moment to change an ending, but the moments soon mount up, affecting the economics and acceptability of MT. One text type for which MT is suitable is minutes of meetings, but French and English write these in different tenses--past in English, present in French--and changing all the verbs is extremely time-consuming. The European Commission has therefore developed an algorithm to change the tenses of minutes automatically between these two languages. It works for almost all tenses, and for all common ones. (It also influences vocabulary, so that, for example, *chair* comes out as *présidence* instead of *chaise*.)

Changes of synonym can also mean extra work for the posteditor. If, in a translation into an inflected language such as French, a noun is changed to a synonym of a different gender, endings in associated adjectives and verbs must also be changed. If the synonym has been offered by ALPnet's interactive MT software (which asks the editor to choose between alternatives), the software will correct the endings. Other systems too may generate the correct inflections automatically once a noun has been changed.

Testing. Testing has determined the level of support for machine translation, whether it is funding for research or implementation in the field. Most spectacular, of course, was the ALPAC report, now increasingly recognized as of dubious quality. In particular, the ALPAC committee assumed that machine translation must not be edited, although much human translation is edited as a matter of course. What it rejected was therefore only Fully Automatic High-Quality Translation (FAHQT), and not, as is often thought, machine translation in general.

To take only three of the other defects, the test passages consisted of sentences taken from six translations (three human and three machine) and jumbled at random, destroying cohesion; the sentences were judged in isolation; and half of the evaluators were Harvard undergraduates instead of real translation users. The ignorance of discourse was general at the time, but the committee's ignorance of translation was less excusable, for it consulted translators little.

In fact, the evaluation of MT is not an easy matter. The Commission of the European Communities, concerned to evaluate its own systems, gathered experts from many countries for a symposium on evaluation in 1978, but no consensus was reached (Van Slype 1979). The Commission's own evaluations had tried numerous criteria (Van Slype 1980). Only two had shown any correlation: postediting ratio (the proportion of words amended) and intelligibility.

My patent feasibility study in 1979/80 was therefore to use these two criteria. Neither, however, was satisfactory. Patents, notoriously, are often imperfectly intelligible in the original. Translations of them may therefore be less than intelligible and yet still accurate, desired, and consequently useful. The postediting ratio was not entirely reliable or suitable, particularly at that early stage in development, and was therefore replaced by an accuracy evaluation.

From my experience of examining professional translators, I felt that the evaluation of translation was inevitably subjective; and that it was better to

acknowledge this subjectivity and concentrate on reducing it. One way to reduce it is to specify the use for which the translation is required. This, as it were, gives a restricted definition of *translation* for the particular evaluation concerned. I therefore asked my evaluators to assume that the text was wanted only for scientific or technical information.

In addition, a further criterion of 'usefulness' evolved. Patents can be translated for several different purposes and to very different standards. Evaluators were asked whether the MT was suitable for other uses than for information.

The usefulness criterion as applied in that study was somewhat crude, but correlated well with accuracy. However, 'usefulness' is probably somewhat too weak a criterion. An evaluator's surprise that the machine can produce anything helpful may make him overenthusiastic, and a stricter criterion such as 'usability' should be applied if the public is not, as in the past, to be disappointed at an early stage.

At the request of the Commission, evaluators were also asked, 'Would the text be useful for postediting?' However, one can postedit to any standard, and the question is unanswerable unless it specifies the purpose for which the postedited translation is required. One evaluator actually suggested that the answer could be obtained by means of the formula:

$$L \times 2$$

2

where L = length of piece of string.

As a translator, I perceived the standard of the MT as very low. My evaluators, however, were translation users (patent attorneys or research chemists), and were more lenient. In a sense both user and translator are right, for users know exactly what they want from a translation, whereas the translator must try to anticipate the needs of numerous, unknown users by supplying a foolproof and 'multivalent' translation. Some users, to save time or money, can accept lower standards than a translator dare supply. Certainly the goal that ALPAC insisted on but rejected--the combination of 'fully automatic' and 'high quality'--is now seen as unrealistic for the present, except perhaps by inexperienced researchers. Instead, human assistance and/or low quality (by our standards) are not only expected, but acceptable, such are the speed and volume which MT can offer.

Even now there is no widespread agreement on how to test. Developers tend to evolve their own mixture of methods. Vasconcellos (1988), after discussing the ALPAC report in detail and surveying other approaches, recommends a mixture of formal and functional criteria.

Perhaps it should be emphasized that MT varies immensely in quality. Firstly, raw MT varies with the inherent suitability of the text for MT, the similarity of the source and target languages, the quality of the system, its experience with the domain and text type, and consequently the size, relevance and depth of coding of its dictionaries. Secondly, postediting varies from rapid postediting of only the most glaring errors to full postediting, possibly to a standard not perceptibly lower than human translation. I once read a whole

paper in Georgetown's *Jerome Quarterly* without recognizing that it was a postedited machine translation (Santangelo 1986).

Even raw machine translation may be suitable for some users. Motivated subject specialists can decode a message even if half of it is lost. A public notice in Cornwall had lost more than half of its characters:

A
AID A D
DI A ED?

I understood it at once. So would you, if you saw it while walking out of a parking lot--provided that you were familiar with our 'pay and display' parking: an open car park, where you buy a ticket from a machine and display it on your car. As a motivated subject specialist, I knew instantly what the notice said.

HAVE YOU
PAID AND
DISPLAYED?

I bought a ticket. We are glad of the redundancy in language when we are not paying full attention, but we do not need it when we are properly motivated--and I did not want to pay a fine.

Technology. Progress in technology and machine translation go hand in hand. The new tool which had brought MT into being was at first slow and crude. Text input was on punched cards, painfully tedious and so expensive in the United States that the cards were punched in Germany and flown across the Atlantic.

Users pushed continually for improvements in quality, speed, and cost. Gradually, matters advanced. Computers became faster, with much larger memories, so that MT dictionaries could be larger and more powerful. Translation was now quicker and better, but input and pre- and postediting were still difficult. Postediting was performed by making handwritten amendments on the printed translation--another very tedious task, which produced a dog's dinner for the typist or customer.

Then came the wordprocessing revolution. Not only are pre- and postediting now far easier, but many source documents are prepared with wordprocessing and can therefore be input without being rekeyed. This cuts both the cost and the time involved in processing a translation. According to one international organization, any human intervention adds at least three days to turnaround time. MT services increasingly, therefore, insist on source texts being submitted in machine-readable form.

I could see from the start that, because typographical and grammatical errors in the source text interfered with analysis, MT needed spell checkers. Soon I heard of the first report-writing program, and knew an automatic pre-editor would follow. This was to be what we now know as text critiquing software. Then the US Air Force developed a semiautomatic posteditor.

Nor were these the only developments to speed the cleaning up of input and output. A more recent tool, the optical scanner, has made another

dramatic difference. At the Nuclear Research Center in Karlsruhe, in Germany, a 500-page translation which could take a month to input now takes only a short time. The translation itself takes only minutes to run, and is then ready: their scientists use raw MT, and so there is no delay for editing.

The next breakthroughs, we are told, will come from artificial intelligence and parallel processing. To date AI has been applied only on a very small scale. MT systems are very large, and it may prove prohibitively expensive to incorporate sufficient AI in them to make a major difference. The greater speed and power offered by parallel processing are attractive and will be useful, but will do little to help with the central task, the analysis of language.

For that, there is no substitute for the process of testing the machine on large quantities of real text, seeing where it fails, and teaching it to do better. That process is a fascinating one. For just as the damaged brain tells us about the healthy brain, the failures of the computer to process language tell us about language. The path pioneered by Georgetown thirty-five years ago is still the way forward.

Practice is the best of all instructors.

References

- ALPAC. 1966. *Languages and machines: Computers in translation and linguistics*; A report by the Automatic Language Processing Advisory Committee. National Research Council Publication 1416. Division of Behavioral Sciences, National Academy of Sciences, National Research Council, Washington, D.C.
- Hutchins, W. John. 1986. Machine translation: Past, present, future. Chichester: Ellis Horwood, distr. Wiley.
- Lawson, Veronica. 1980. Final report on EEC study contract TH-21. Feasibility study on the applicability of the Systran system of computer-aided translation to patent texts. CETIL/205/80. Luxembourg: Commission of the European Communities.
- Lawson, Veronica. 1982. Machine translation and people. Practical experience of machine translation. Proceedings of a conference, London, 5/6 November 1981, ed. Veronica Lawson. Amsterdam: North-Holland. 3-9.
- Morgan-Girard, Danièle. 1981. Personal communication.
- Santangelo, Susana. 1986. Machine translation: Where to? / Traducción automática: Rumbos. Jerome Quarterly 1.4:4-5.
- Van Slype, Georges. 1979. Critical study of methods for evaluating the quality of machine translation. Final Report. Brussels, Luxembourg: Commission of the European Communities.
- Van Slype, Georges. 1980. Systran: Evaluation of the 1978 version of the Systran English-French system of the Commission of the European Communities. Incorporated Linguist 18.3:6-69.
- Vasconcellos, Muriel. 1988. Factors in the evaluation of MT: Formal vs. functional approaches. In: Technology as translation strategy, ed. Muriel Vasconcellos. Binghamton: SUNY. 203-13.

Acquisition versus learning in reading pronunciation by adult EFL students

Robert Lado

Georgetown University

School of Languages and Linguistics, Dean Emeritus

This pilot study seeks to generate objective data on a major assumption, sometimes explicit, sometimes implicit, of most of the second language teaching methods proposed in the 20th century: that second language learning by students and first language acquisition by children are or ought to be the same. Since all nonimpaired children acquire their native language seemingly without effort and without formal teaching, it is further assumed that there must be a simple key to the language acquisition process. And arbitrary restrictions are imposed on adult second language learning to fit particular models of first language acquisition.

Each new method presumes to have found the simple key, but the proposed solutions contradict each other, and none produces the uniform, easy, and rapid acquisition, or native pronunciation, of the second language taken for granted in first language acquisition.

The experiment is also intended to be a partial test of Krashen's acquisition-learning hypothesis (1982), which claims that adults acquire a second language via the same language acquisition device as children.

The acquisition-learning distinction is perhaps the most fundamental of all the hypotheses to be presented here. It states that adults have two distinct and independent ways of developing competence in a second language.

The first way is language 'acquisition', a process similar, if not identical, to the way children develop ability in their first language...

The second way to develop competence in a second language is by language 'learning'. We will use the term *learning* henceforth to refer to conscious knowledge of a second language, knowing the rules, being aware of them, and being able to talk about them...

Some second language theorists have assumed that children acquire, while adults can only learn. The acquisition-learning hypothesis claims, however, that adults also acquire, that the ability to 'pick-up' languages does not disappear at puberty. This does not mean that adults will always be able to achieve native-like levels in a second language. It does mean that adults can access the same natural 'language acquisition device' that children use. As we shall see later, acquisition is a very powerful process in the adult...

Evidence from child language acquisition confirms that error correction does not influence acquisition to any great extent... (Krashen, 1982:10-11)

Krashen and Terrell restrict language acquisition to Krashen's Input Hypothesis:

The central hypothesis of the theory is that language acquisition occurs only in one way: by understanding messages. We acquire language when we obtain comprehensible input, when we understand what we hear or read in another language. (Krashen and Terrell 1983:1)

Krashen postulates further that learning experience does not contribute directly to acquisition, that what is learned remains conscious learning and does not become acquired competence.

Acquisition occurs, according to the input hypothesis, when acquirers understand input for its meaning, not when they produce output and focus on form. (Krashen 1982:117)

A different view is that when a phoneme, word, pattern, or rule is not readily assimilated in the learner's symbolic language and discourse systems by simple acquisition, it is assimilated through learning and use and eventually becomes part of the speaker's competence.

Chomsky's LAD. The easy, child language acquisition assumption received strong support from Chomsky's postulation of an organlike language acquisition device (LAD) which innately possesses syntactic structure and is programmed to achieve native language competence from imperfect samples of the language heard by children. Krashen's reference to the same 'language acquisition device' is to Chomsky's LAD (Chomsky 1965, 1980). For a discussion of the LAD in the context of second language acquisition and learning, see Lado 1985.)

Objective. The immediate objective of the experiment is to compare, experimentally, the effect of acquisition experience with teaching-learning experience in reading pronunciation¹ by adult EFL students.

Hypothesis. The principal hypothesis is that the reading pronunciation of the phoneme /δ/ as in *this, other, bathe*, one of the most difficult to master by EFL students of several language backgrounds, will improve significantly more with specific teaching-learning intervention than with general acquisition experience in English classes. The hypothesis is disproved by failure to find any difference between improvement in the reading pronunciation of /δ/ through teaching-learning experience and improvement in the pronunciation of all other sounds through acquisition experience. The hypothesis is also disproved if no difference is found between improvement in the pronunciation of /δ/ with teaching intervention and without it.

Design. A test-teach-test design for the experimental samples was used. The subjects' reading of the selection, *Grip the Rat*² was recorded on cassettes. Subjects then received three or four³ 20-minute lessons on /δ/ on successive days, and recordings were made of their reading of the same selection after the last lesson. The selection itself was not taught. Performance on the pronunciation of /δ/ and separately on the other sounds was scored and compared.

Improvement in the pronunciation of /δ/, which was explicitly taught, is ascribed to learning. Improvement in the other sounds is ascribed to acquisition since there was no teaching of the other sounds--subjects merely listened to and read aloud the material of the /δ/ lessons with comprehension.

Subjects. The subjects were literate adult EFL students of different native language backgrounds at levels 4 and 5 of a 10-level intensive course. The course meets four hours a day five days a week in groups not to exceed ten students per group. Level 10 subjects are expected to score better than 500 on the TOEFL. Levels 4 and 5 are intermediate. Students are placed in those levels on the basis of their average score on written and listening placement tests.

Scoring. The cassette recordings were listened to as many times as necessary on a variable speech recorder to determine whether the sound was the phoneme /δ/ or some other, usually /d/ but sometimes /θ/, or /z/. All other sounds were scored by similar criteria and whether or not the distortion was sufficiently noticeable to constitute communicative noise. Intrusive sounds such as intrusive initial /h/ by French speakers and intrusive initial /e/ by Spanish speakers were marked wrong. Initial /r/ pronounced by Portuguese speakers as a voiced velar fricative is perceived as a strong /h/ in English (*rat* sounds like *hat*) and was marked wrong. Stress and pause distortions that incorrectly changed the particle of a phrasal verb into the preposition of a following prepositional phrase were marked wrong. Tag questions with the clear rising intonation of a full question were also marked wrong.

Every occurrence of /δ/ was counted as either right or wrong. The score was the percentage of correct pronunciations. For the other sounds, the scoring was on the basis of words pronounced correctly. One mispronunciation in a word made the word wrong. Most occurrences of /δ/ were in the word *the*. In addition to counting the right and wrong occurrences of /δ/, the word itself was tallied right or wrong for other sounds on the basis of the vowel. When the vowel of *the* was pronounced /e/ instead of /ə/ before a following initial consonant, or /e/ or /ə/ before a following initial vowel sound, the word was counted wrong. There are 539 words and 68 occurrences of /δ/ in the selection.

A number of subjects took only the pre- or the posttest and could not be used for comparison of improved pronunciation. Students who missed one of the four lessons were included in the tabulation.

One subject with a perfect score on /δ/ before the lessons was excluded on the ground that there was no room for improvement. Eleven subjects in

a fourth group had to be eliminated because the teacher taught the sound /δ/ and practiced it in preparation for the experiment.

Lessons. Written handouts of the lessons were distributed to the subjects before each lesson. The teaching-learning experiences included study of facial diagrams showing the articulation of /δ/, /d/, and /θ/, and articulatory description of the three sounds, noting their contrasting as well as their similar features. The lessons included practice with minimally contrasting pairs of words and similar pairs of words and phrases in initial, medial, and final positions, and the pronunciation of the vowel of *the* as /ə/ before words beginning with a consonant sound and /i/ before words beginning with a vowel sound. The rule was made explicit in connection with practice.

Sentences with /δə/ and /δi/ were practiced. Phrases with *this* and words beginning with *dis-* were contrasted for sound and meaning, for example, *display: this play, distribute: this tribute*. Proverbs with examples of /δ/ were recited. Two fables, *The Lion and the Statue* and *The North Wind and the Sun*, were explained and recited, as were the first fourteen lines of *Mending Wall* by Robert Frost, and all of *Annabel Lee* by Edgar Allan Poe.

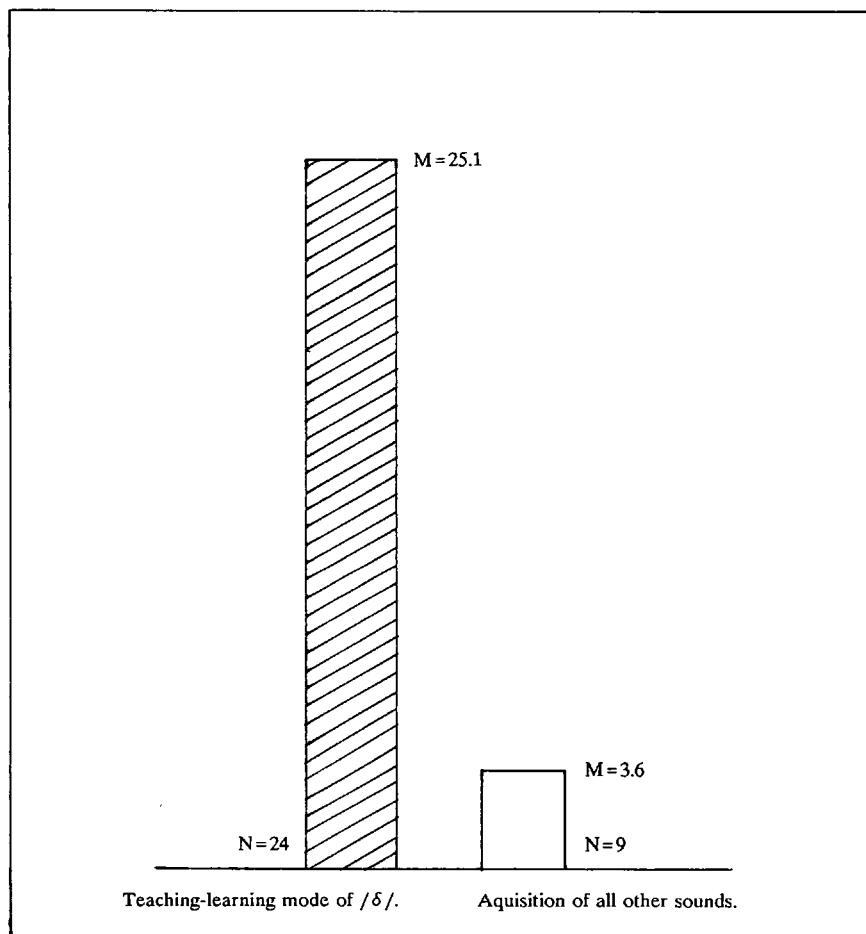
Results. The results show an improvement of 25 percentage points in the reading pronunciation of /δ/, and 3.6 percentage points in the pronunciation of all other sounds (Table 1 and Figure 1).

Table 1. Improvement in percentage scores by adult EFL students in reading pronunciation of /δ/ with teaching and in all other sounds without. Reading selection: *Grip the Rat*.

Readings: Student	/δ/:			All other sounds:		
	1st	2nd	Differ- ence:	1st	2nd	Differ- ence
1.F.F.	32	55	23	80	80	0
2.S.P.	22	85	63	75	81	6
3.A.C.	6	16	10	65	68	3
4.A.Z.	36	56	20	71	73	2
5.R.G.	45	94	49	63	65	2
6.C.W.	20	75	55	45	48	3
7.S.S.	44	65	21	64	78	14
8.T.N.	21	49	28	48	51	3
9.T.C.	33	68	35	88	89	1
10.C.S.	90	87	-3	82	84	2
11.C.A.	35	43	8	82	90	8
12.M.D.	31	78	47	81	92	11
13.J.B.	75	84	9	84	85	1
14.R.E.	27	79	52	78	87	9
15.F.V.	44	49	5	78	81	3
16.A.N.	31	79	48	81	92	11
17.I.J.	40	58	18	84	88	4

18.N.A.	43	77	34	83	85	2
19.A.B.	80	96	16	86	87	1
20.C.H.	71	96	25	90	91	1
21.J.B.	39	52	13	57	57	0
22.D.L.	44	64	20	83	79	-4
23.V.G.	66	71	5	88	88	0
24.M.Q.	45	44	-1	72	75	3
N=24	$\Sigma = 1020$	1620	600	1808	1894	86
	M= 42.5		67.5	25	75.3	78.9
						3.6

Figure 1. Improvement in reading pronunciation of /δ/ in teaching-learning mode vs. improvement in all other sounds in acquisition mode.



All the sounds were heard, pronounced, and experienced meaningfully in an acquisition mode. Only /δ/ was studied and practiced consciously trying to learn it in a teaching-learning mode. The LAD acquisition hypothesis and methods based on it assume that conscious teaching and practice are useless, that learning has no effect on acquisition, and that students will acquire the language (which includes pronunciation) merely by experiencing it as communication.⁴

Reading aloud *Grip the Rat* a second time after the lessons on /δ/ could presumably allow Krashen's Monitor to intervene, but the reader is recording the whole selection, not just the words with /δ/, and is trying to read it meaningfully as shown by intonation, stress, and pauses. The reader has no time to apply rules of pronunciation, and the conditions for Monitor correction are not present.

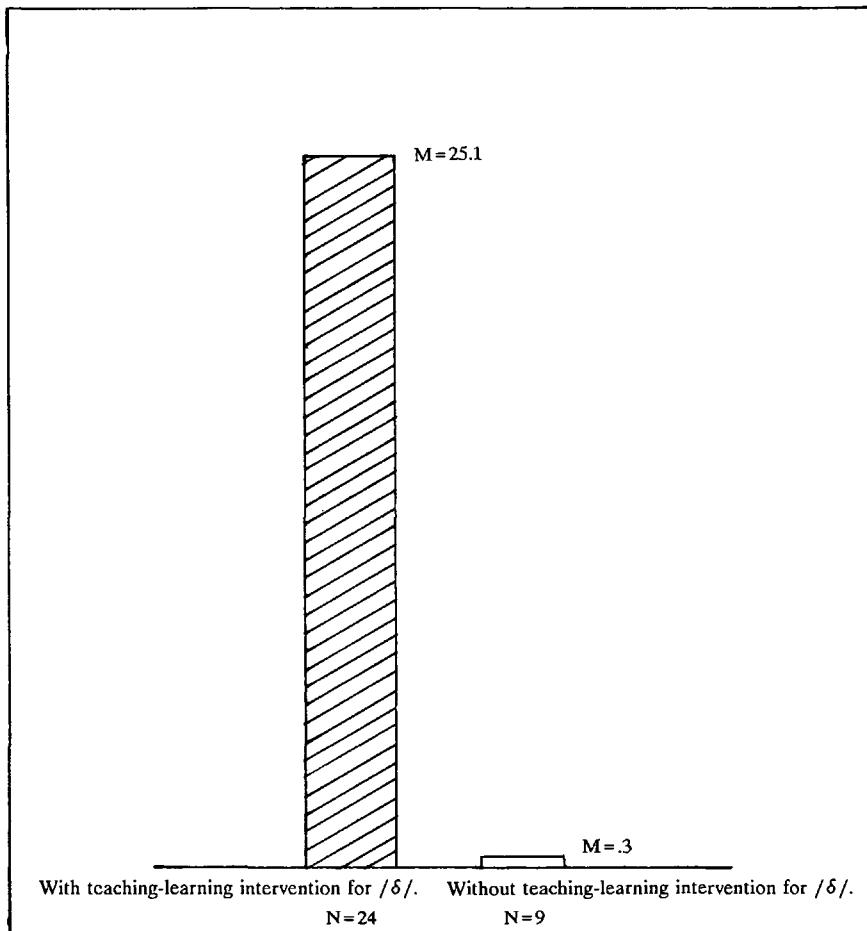
Two additional questions were asked in the experiment. (1) Will students who do not receive specific teaching in /δ/ in the same intensive course improve in its pronunciation in the same period of class days assigned to the lessons? And (2) will the improvement in /δ/ last, or will it show the steep curve of forgetting typical of studied material?

To answer the first question, I had ten subjects at level 5 record the selection four days apart without intervening lessons on /δ/; they followed the regular lessons of the same intensive course. One subject did not appear for the second reading because of illness. The results with the nine subjects that recorded both readings of *Grip the Rat* were as follows: Mean percent of correct pronunciation of /δ/: first reading, 41.9; second reading, 42.2. Improvement in mean percentage points, .3. This contrasts sharply with the improvement of 25 following specific instruction on /δ/ (Table 2 and Figure 2).

Table 2. Improvement in reading pronunciation of /δ/ by adult EFL students without teaching.

Student	1st Reading	2nd Reading	Improvement
1. G.T.	12	24	12
2. S.H.	40	47	7
3. A.C.	50	47	-3
4. C.N.	64	62	-2
5. M.F.F.	60	57	-3
6. S.A.	54	34	-20
7. N.O.	40	31	-9
8. A.T.	34	32	-2
9. S.G.	23	46	23
N = 9	$\Sigma = 377$	380	3
	M = 41.9	42.2	.3

Figure 2. Improvement in reading pronunciation of /δ/ with teaching-learning intervention and without.



To answer the second question, whether or not any improvement in the pronunciation of /δ/ would be temporary and show the steep curve of forgetting typical of studied material, three subjects were recorded a third time later after the lessons. The results do not show the steep curve of forgetting typical of learned material. They show, instead, no loss for a period of 125 days during which the subjects continued studying English intensively four hours a day five days a week without explicit teaching of /δ/. If acquisition had been effective in improving their pronunciation of /δ/, which they heard thousands of times pronounced by their native English teachers, some improvement over 125 days should have appeared. In other words, the

subjects maintained what they had learned and did not acquire any improvement in the reading pronunciation of /δ/.

The same set of lessons on /δ/ were repeated once again in 20-minute sessions with these three students, and their reading of *Grip the Rat* was recorded a fourth time following the lessons. Their reading pronunciation of /δ/ increased to 83%, an improvement of 47 percentage score points over the 36% they scored before the second round of teaching.

If additional subjects show the same spurt of progress following teaching-learning intervention and extended flat plateaus when the sound is not taught but merely experienced in intensive acquisition experiences, we could conclude that fossilization of the mispronunciation will set in unless teaching-learning experience on /δ/ is provided (Table 3 and Figure 3).

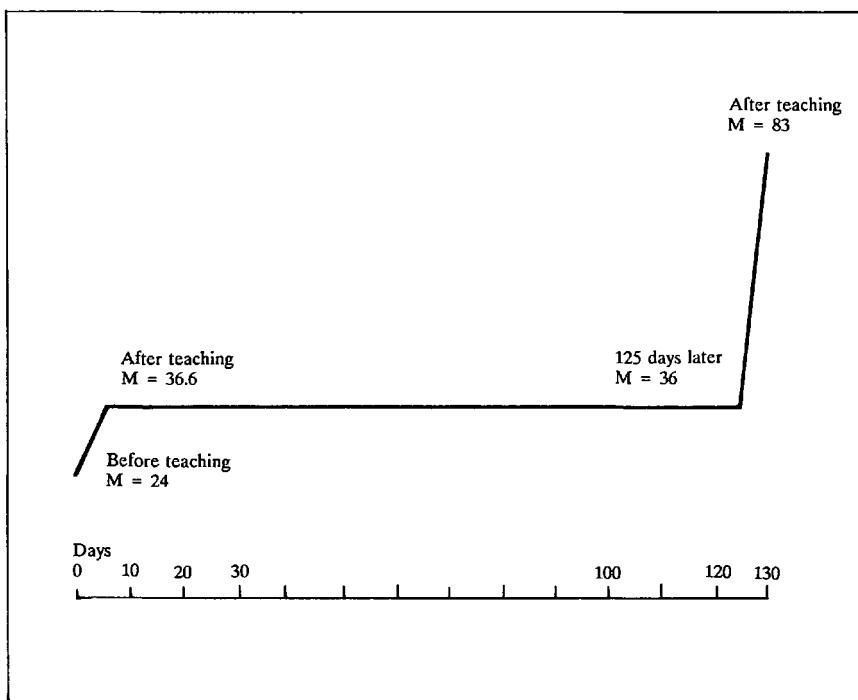
Table 3. Improvement in reading pronunciation of /δ/ before teaching, after explicit teaching of /δ/, after 125 days studying English intensively with native English teachers in the U.S. without further teaching of /δ/, and after a second round of lessons on /δ/.

Student	1st Reading	2nd Reading	Delayed	3rd Reading	4th Reading
1. F.F.	32	57		51	63
2. C.A.	35	43		38	93
3. A.C.	5	10		19	93
N=3	$\Sigma = 72$	110		108	249
	M = 24	36.6		36	83

Conclusions and recommendations. Acquisition experience alone is less effective for the achievement of good reading pronunciation by adult EFL students than teaching-learning experience. Many adult EFL students who need to read aloud are motivated to attain good or even nativelike pronunciation if that were possible. Such a goal seems more likely to be achieved with teaching-learning techniques and oral reading aided by adequate recording equipment than with acquisition experience alone.

Implications. If these findings are confirmed by further experimental evidence, we can speculate that after a century of searching without success for a simple key to achieve second language competence on a model of easy first language acquisition by children, it is time to admit that the achievement of second language competence by adults shows major differences with first language acquisition by children. Adults think differently, remember differently, understand differently. They know more than children, have acquired and learned ways of doing things that facilitate and retard the acquisition and learning of new things. They have different motivations, less time and more inhibitions, and have other interests, drives, and choices.

Figure 3. Reading pronunciation of /δ/ before teaching, after teaching, 125 days later, and after teaching /δ/ again.



I propose that at least some of us break away from the fixation of the first language acquisition model and focus attention on theory, research, and teaching that relate more directly to our real students, in real environments, taught by real teachers, with real travel and technological resources.

Notes

1. Reading pronunciation is not a free skill, it is part of reading aloud; it is bound to reading as pronunciation is bound to speaking, and without reading pronunciation there can be no reading aloud, as without pronunciation there can be no speaking. Reading pronunciation depends on language competence, general reading ability, and knowledge of content.
2. I listened to recordings of *Grip the Rat* showing differences in geographic dialects of English but have thus far been unable to find its original source. The version I used was transcribed into conventional spelling by me from the phonetic transcription of Nelson Francis' own speech (Francis 1958: 116-18).

3. Subjects who were absent from class on one of the four lesson days were included in the results.

4. Terrell (1989) seems to have broken away from the strong comprehensible input acquisition hypothesis by redefining the role of grammar in a communicative approach. He suggests examples of teaching explicit grammar in preparation for what he terms 'binding/access activities' that resemble my 'pattern practice' in context. 'The central idea of "binding/access activities" is to limit the grammatical choice in focus to a single meaning-form relationship making it so frequent and salient at the same time that the students will focus on meaning and form,...' He also upgrades the role of output with an obscure suggestion that 'In addition I suggested that monitoring might affect acquisition if it is possible for learners to acquire their own output.'

Terrell justifies these changes because 'an instructor can give good comprehensible input using Spanish past tense forms on numerous occasions without the students even attending to or acquiring Spanish past tense morphology at all' (p. 24).

He proposes, for example, that with gender endings in Spanish, besides the possibility of waiting for the acquisition process to run its natural course, 'it is my [Terrell's] view that this is not necessarily the most efficient approach in the classroom and perhaps also not as efficient even in a natural second language context.' The other [possibility] is to use interventionist techniques,... One possible focus for intervention techniques is explicit grammar instruction' (p. 23).

Terrell also makes use of contrastive explanations based on transfer when he talks of English speaking students not being 'accustomed to attending to a determiner to signal plurality' in French.

References

- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.
 ———. 1980. In: *Language and learning, the debate between Jean Piaget and Noam Chomsky*, ed. Massimo Piatelli-Palmarini. Cambridge, Mass.: Harvard University Press.
 Francis, W. Nelson. 1958. *The structure of American English*. New York: The Ronald Press Company.
 Krashen, Stephen D. 1977. The Monitor Model for adult second language performance. In: *Viewpoints on English as a second language*, ed. Marina Burt, Heidi Dulay, and Mary Finocchiaro. New York: Regents.
 ———. 1982. *Principles and practice in second language acquisition*. Oxford: Pergamon Press.
 Krashen, Stephen D., and Tracy D. Terrell. 1983. *The natural approach: Language acquisition in the classroom*. Hayward, Calif.: The Alemany Press.
 Lado, Robert. 1985. 'Total' approach to second language teaching and learning. In: *Current approaches to second language acquisition: Proceedings of the 1984 University of Wisconsin-Milwaukee Linguistics Symposium*, ed. Barbara Wheatley, Ashley Hastings, Fred R. Eckman, Lawrence Bell, Gary Krubar, and Rita Ruthkrowski. Reproduced by the Indiana University Linguistics Club, 310 Lindley Hall, Bloomington, Indiana 47405. 110-126.
 ———. 1988. *Teaching English across cultures*. New York: McGraw-Hill.
 Perdue, Clive, ed. 1984. *Second language acquisition by adult immigrants: A field manual*. The European Science Foundation Project. Rowley, Mass.: Newbury House.
 Terrell, Tracy David. 1989. The role of grammar in a communicative approach. *Northeast Conference on the Teaching of Foreign Languages*. Newsletter No. 25, February 1989. 22-25.

Discourse and text: A narrative view of the foreign language lesson

Claire Kramsch

Massachusetts Institute of Technology

The setting is a fifth grade English class in Germany. The lesson plan calls for vocabulary explanations prior to the reading of a text by the class. The teacher shows the picture of a boy and his dog and asks:

Excerpt 1:

... and does anybody know how we call this part of the dog? Can you see it? This part.

S. Tail

T. Yes, it's the tail of the dog. And what is Toby doing, doing with his tail? Can you see this?

S. He is waving with his tail.

T. Well, he isn't waving--he is--well the correct word is 'he is wagging his tail'. (writes on BB: to wag his tail) So who can give me a definition of 'to wag a tail'? What does a dog do if he wags a tail?--wags his tail.

S. He is happy.

T. Yes, he is happy, and you said the reason why he is doing this. But who can give me a definition of the word 'to wag'? What does this mean 'to wag a tail'?

Ss. (silence)

T. What does a dog do if he wags his tail, if he wags it?

S. He likes to play with you.

T. Yes, he likes to play with you. Well, he is moving the tail from one side to the other side. This means 'to wag a tail'. (Nold, 1983)

Form vs. function: A dichotomizing stereotype. What's going on? One could say that the teacher is trying to elicit orally from the students a written, dictionary definition of the expression 'to wag a tail' which he can then write on the blackboard and the students in their notebooks, while the students are, consciously or unconsciously, insisting on 'meaningful' responses. The teacher is trying to decontextualize the phrase 'to wag a tail', requesting a universal context-free response--what Watson and Olson have called the 'literate bias on the structure of word meaning' (Watson and Olson 1987), whereas the students insist on giving particularistic, context-specific interpretations. In a way, the students are responding to the teacher as if he were a real-life person in a real-life communicative exchange, while the teacher tries to focus their attention on classroom academics.

Exchanges such as this one have been used recently in classroom research to illustrate 'communicative deficits' of teacher behavior in language classrooms: lack of communicative information, gap and undue attention to form over message (Fröhlich, Spada, and Allen 1985), excessive number of display questions (Long and Sato 1983), disproportionate amount of teacher talk and management of teacher talk vs. learner talk and learner control (Kramsch 1984), procedural navigation instead of negotiation of meaning (Allwright 1980), lack of coherence in classroom discourse (Nold 1983). Already the analysts' choice of these negative terms implies that teacher behavior needs to be 'redressed' and communication in the classroom made more 'naturally' interactive (Rivers 1975).

With the push for 'authentic interaction' and communication, and the development of oral proficiency in real-life contexts, teachers feel increasingly caught in the dichotomy between the forms and functions of language in use. Especially in educational settings, the question that is always asked is still: how much focus should be given to form, i.e. to language as *text* fixed by norms of structure and usage, how much to function and content, i.e. to language as *discourse* that varies with situations and contexts of communication?

I will argue here that the form/function dichotomy is a stereotype of our educational culture and the result of a too narrow linguistic view of communication. I will try to rephrase the problem by proposing instead a narrative view of foreign language teaching that draws on insights gained in ethnography and social psychology.

Discourse vs. text. Let us go back to our original exchange. What went wrong? Is it the fact that the teacher was operating in the written text mode of the dictionary, the students in the spoken discourse mode?

Spoken vs. written language. Due to the current emphasis on oral proficiency and the recognition of the discourse component in communicative competence, foreign language teachers often equate discourse with the spoken, text with the written language: students engage in oral communication activities and make use of sociolinguistically appropriate gambits in *discourse*, especially at the elementary levels; they read, write, classify and draw information from *texts*, especially at the intermediate and advanced levels.

And yet, as has been amply shown by Deborah Tannen (1984) and others, the spoken/written distinction is not as clear cut as it seems. For example, in the spoken exchange cited above, the teacher is trying to elicit orally from the students a textual dictionary definition of the expression *to wag a tail*, which he can then WRITE on the blackboard, while students' responses are typical of oral conversation. Rather than spoken vs. written, we should perhaps use Tannen's distinction between oral and literate forms of discourse to examine foreign language lessons. Tannen points out that it is not 'spoken' vs. 'written' per se that is at issue here, but rather 'the relative focus on communicator/audience interaction' (which she calls 'interpersonal involvement') on the one hand, as opposed to the relative focus on 'message content' on the other (1984:3).

Focus on interaction vs. focus on message content. Concern for the discourse component of communication has focused attention on classroom interaction. Teaching a foreign language is seen as teaching students how to interact with written texts and with interlocutors in various communicative situations (Rivers 1987). Students need as much input (Krashen 1982) and as many opportunities to negotiate the meaning of that input as possible (Swain 1985) under the control or even without the control of the teacher. The focus is on how to do things and how to have others do things with words in concrete situations.

However, for all its flexibility to various interlocutors and situational contexts, this focus on discourse and interaction is constrained by an equally necessary focus on message content, and on the appropriate norms of social behavior. For if students were always engaged in small talk in personalized situational contexts, even if in the foreign language, they would never be schooled into an academic educational world that requires among other things the ability to abstract and generalize and to pay attention to linguistic forms, cultural concepts, and discourse topics. Such decontextualized knowledge is, as we know, the source of great power.

Indeed, much of what we have to teach has to do with the message content of language: rules of grammar and vocabulary (including dictionary definitions), retrieval of information from spoken or written documents, collection, transmission and organization of this information in the form of grammatical, lexical, cultural or narrative facts, and group activities based on exchange of information among participants. The focus is not only on what to do, but what to say and how to say it with words (Kramsch 1988).

Discussion. However, here again, the pedagogical boundaries between focus on interaction and focus on message content are not clear cut. Nicole Soule-Susbielles, in a recent article in *Langues Modernes* (1988), describes a paired communicative activity ('Tell each other your life story') which one of the students 'performed' as an imposed assignment with no interest whatsoever in his interlocutor, whereas she, by contrast, took it as an opportunity to really get to know a classmate. Is focus on content and focus on interaction a matter of perception and intention? Similarly, a spontaneous, 'personalized' dialogue between teacher and student, but in which the teacher is primarily interested in making the student 'talk', is quickly perceived by the student as an academic exercise in message-transfer. Even 'informal conversations', if measured against preestablished norms of gambitry, can be more trouble-shooting than real negotiation for meaning (Aston 1986). By contrast, a spoken or written message content can be the object of much interpersonal involvement if the focus is not on any textual norm, but on the negotiation of relative and varying meanings between speakers and hearers, between readers and written texts, and on the different interpretations given and received.

Thus it seems that the concepts *discourse* and *text* in the language classroom do not denote different bodies of knowledge nor do they apply to different activities: here interpersonal communication, there textual analysis; here functional proficiency, there attention to linguistic accuracy and truthfulness of message. Nor can one say that the former is more

characteristic of elementary levels of learning with its emphasis on functional tasks and communicative doing, whereas the latter would be more appropriate for the intermediate and advanced levels with their emphasis on elocution and style.

In fact, discourse and text, as Tannen has described, indicate two inseparable and complementary relative foci on the oral and literate modes of communication at all levels. For foreign language teaching, however, I would like to replace her idea of an oral-literate continuum by another metaphor that could better describe the communication-oriented language classroom. I would like to view the foreign language lesson as a conversational narrative that is jointly told by teacher and students for each other's benefit. This narrative consists of a script and its enactment, or, in Hymes' ethnographic terms, of a text and its performance within a narrative view of life (Cazden and Hymes 1978).

Let us examine how this metaphor of the conversational narrative could help us out of the form/content dichotomy and of a traditionally asymmetrical teacher/student power relationship that is difficult to harmonize with the development of symmetrical communicative competence.

Text and performance: A narrative view of the foreign language lesson.

Telling the foreign language 'story'. Every narrative in a conversation is both a *performance*, carried out through the interaction of teller and hearers, and of potential tellers among the hearers, and a *text*, that is all the more pleasurable to the participants as its structure is predictable and its content has familiar echoes. One can view the language lesson as a narrative, told through the joint efforts of teacher and learners, who try and build common topics across cultural and social differences via the foreign language. The text of this narrative may seem predictable, given by textbook, syllabus and the expectations of schooling, but its performance, ever renewed with every new group of learners, gives it a nonreproducible life that is more than the sum of its parts. This narrative is made of many individual 'stories', each with its unique voice, its unique style, its unique viewpoint. The term 'conversation' is taken here to denote the dialogic frame that defines the speech community of the classroom within one or a series of language lessons. Adopting a Vygotskyian perspective, I consider that students learn both the foreign structures and their use through dialogue and by internalizing patterns of conversation or interaction into patterns of thought and speech (Vygotsky 1962).

Now the *raison d'être* of a narrative is in its telling: staging or *mise en scène*, tellability and relevancy, as well as a certain balance in the turn-taking and the rhythm of narration, are key factors in its effectiveness. The best narrative in conversation can be frustrating if it is not timed and framed properly, or if the listeners don't 'see the point', i.e., don't understand the relevancy of the story to what they perceive the general context or topic of conversation to be (Polanyi 1979)--in other words, if they are not brought into the story that is being told. Prefacing, framing operations, as well as digressions, evaluations, and metacommments, allow its joint construction by

speakers and potential speakers such as listeners, eavesdroppers, and bystanders.

The metaphor of the narrative enables us to see teacher and students not enacting different instructional roles 'as actors learning the lines from some pre-scripted target repertoire for a performance at some later time and place' (Breen and Candlin 1980:99), but engaged in narrating a story that has its own here-and-now circuit of meaning, its own flashbacks, its own network of internal and intercultural references.

Let us consider the following excerpt from a second-year English high school class in Switzerland from a narrative point of view (1). The teacher is reviewing the vocabulary encountered in a ghost story that he has just read with the students and is brainstorming words that he writes on the board:

Excerpt 2:

S2. a mirror

T. a mirror (writes on blackboard) it's michael jackson [again? why?
S2. no

T. why?

S2. because a vampire can't see in a mirror

T. I see, he can't see himself in a mirror

S2. no you it's you can't see a vampire in the mirror

T. Oh that's why I haven't seen a vampire this week (everyone laughs) I
see...

The students enjoy listening both to each others' far-fetched ideas and to the spontaneous reactions from the teacher, as he writes every word on the board. Each contribution from the students is commented upon, embedded so to speak in ministories that enrich the main storyline, which the teacher is responsible for maintaining. Then rather abruptly, the teacher turns to an exercise on relative clauses, using the words on the board. The students are immediately turned off.

T. ok fine, can you please write down now--this is a a bit difficult--a definition ; d'you know what a definition means?

Ss. (sigh)

T. definition means

S. (under his breath) awful

T. a description--or can you explain what the word means with a relative pronoun--of course our old friend the relative pronouns--I want you to explain the following--you write down what is a *haunted house*-- explain what that is in English in a sentence--that's a house and so on, or a haunted house is a house...write a definition like in the dictionary - ... alright I give you a second definition a second word, the second word is the word a *cellar*--what is a cellar?

Sensing that the community involvement has been disrupted by this sudden but necessary turn to grammar, the teacher brings students into the story by referring to previous story events in which they were included.

- T. the word *cellar* - who said the word *cellar*, that was your word wasn't it?
S. yes

But the students still resent the formal, seemingly irrelevant exercise on relative pronouns. They recontextualize the story by drawing on the immediate physical environment and by using the native Swiss dialect of that environment.

- T. what's a cellar, what's a cellar, how would you explain it in English simply
(There is a noise of pupils outside in the yard)
S. das isch a geischt gsi (tr. 'that was a ghost')
(pupils laugh)
T. that's a ghost (pupils laugh) ja good but in English now this is the English lesson

To reinstate the sense of community, but to continue with the text he had planned for the lesson story, the teacher then transforms an individual normative exercise into a competitive, polyphonic verbal display.

- T. we can get a lot of definitions, let's see who's got the nearest definition which is exact...
S1. that's a house where ghosts live
T. yes that's isn't a relative clause but it's--that's a house where ghosts live would be an answer...
S2. it's a house where ghosts appear or where you can hear footsteps or chain clangs or any other suspicious noisy and lights...
T. that's a longer one isn't it...
S3. I've got a house which has got a lot of ghosts in it
T. well that's a relative clause which has got a lot of ghosts in it
(students laugh).

From this excerpt we can see that the main task of teachers and students is the joint generation and performance of a text that focuses alternately on the message and on the language used to convey that message; in addition, there is constant evaluation by the participants of both performance and text. These three foci correspond roughly to the three elements that discourse analysts have identified in conversational narratives: events, descriptions, evaluations (Polanyi 1979).

The text

(1) **Events.** Events in the foreign language lesson are the transmission and exchange of new information (grammatical, lexical, social, cultural information) in small groups or with the whole class. The truth of these events is in their reference to the outside world of lexical definitions, grammatical norms, physical properties of ghosts and other cultural phenomena. Their meaning is in the doing. In the world of events, teacher

and students are storytelling partners in a (simulated) foreign cultural situation where people do things with words.

(2) Descriptions. The foreign language class does not consist only of events; like any narrative, it contains long stretches of description. Through display questions, choral repetitions, echo responses, grammatical practices and rituals, participants in foreign language lessons often speak for the sake of speaking and of hearing each other speak. The meaning of these narrative descriptions is in their saying, their truth is in the enunciation itself. Drama techniques, jazz chants and recitation of poems, but also reading aloud, thinking aloud, discussing choice of words and grammatical structures, appropriate registers and pragmatic behaviors (both schooled and natural)--all these expressive and self-conscious uses of language are various ways of 'exhibiting meaning within the language action itself' (McNeill 1985:268). In the world of descriptions, teacher and students are personae in self-created narrative relationships.

The performance. Events and descriptions are performed according to more or less schooled rules of interaction, orchestrated by the teacher, either with the whole class or with pair or group work. We can identify in the excerpts above several aspects of this performance.

(1) Openings or prefaces. Openings such as: 'can you explain what the word means' 'and does anybody know how we call...' set the topic of the vocabulary story the teacher wants to tell; the use of 'we' invites the students to collaborate in the joint telling of that story.

(2) Orientation. Orientation can be social ('does anybody know...'), spatial ('this part'), physical ('can you see it?'), or rhetorical ('d'you know what a definition means'). Sacks argues that 'the strategic use of such contextual information keeps hearers attentive to how to interpret what is being told' (cited in Young 1982:298). In this sense, it seems that the orientation devices used in the second excerpt were more effective than those used in the first, because they were both more explicit of what the teacher wanted, and they allowed more room for students' (dissenting) voice.

(3) Relevancy. Following an assigned reading about ghosts, the teacher recapitulates the narrative by brainstorming vocabulary. This sets up certain expectations in the students who expect to find the point of the activity in its relevancy to the ghost story. As storytelling analyst Katherine Young notes: 'To miss the point is either not to see how the events in it connect together or not to see how they are relevant to the occasion of their telling' (p. 301). If students do not see how a dictionary definition of 'a haunted house' is relevant to understanding or discussing the story, they will 'miss the point' of the teacher's questions or try to create a diversion. This 'point' is lodged in the series of events that preceded the vocabulary activity (e.g. the ghost story), but also in those that will follow it, i.e. the general objectives of the course as stated explicitly by the teacher (e.g. development of communicative competence) or implicitly by the school system (e.g. grammar and vocabulary).

tests, grades etc.). The abrupt switch to the formal grammatical exercise might cast retrospective doubts in the minds of the students about just how genuine the narrative commitment of the teacher had been in the previous brainstorming activity. Most misunderstandings on the part of learners arise because they miss the point or they believe the point is elsewhere.

(4) **Tellability.** Excerpt 3 (see Appendix) shows the breakdown in communication that can occur when teacher and student are trying implicitly to tell two different stories. In this passage, the teacher is gathering arguments for the abstract *topic*: private vs. state medical care, whereas the student is talking about *people* who shouldn't go to the doctor if they only have a cold (as one of the students obviously has). The teacher usually has the power to decide what is tellable in class ('I'm afraid I don't understand'). However, the success of a communicatively oriented lesson will depend on the ability of the several storytellers to understand, accept, and integrate each other's stories--both the teacher's abstract story and the students' contextualized one. It will need an awareness of the fact that any incident in the lesson is potentially of interest and worth telling.

(5) **Presentation of self, not only of knowledge.** A narrative view of the language class makes teacher and students into storytelling persons, rather than conveyors and displayers of disembodied knowledge. Equally important as the way in which knowledge is displayed is the way teacher and students *present* themselves acoustically, visually and linguistically to one another via the foreign language. Variations in the presentation of knowledge (e.g. for some students in a more oral, for some in a more literate mode) are viewed as variations in performance style as they are encountered in real-life communicative situations.

(6) **Building a speech community.** In a recent paper presented at the American and British Associations of Applied Linguistics, Courtney Cazden reports on the relative contributions made by Hymes and Bakhtin to the concept of communicative competence. 'Hymes', she says, 'views "community as an organization of diversity", in which co-existence of language and discourse varieties is peaceful, unless access to the conditions necessary for their acquisition has been denied' (Cazden 1988). Thus, in the foreign language classroom, it would be enough that every student's story be accepted and integrated into the common narrative to make learning successful. By contrast, Bakhtin stresses the 'intra-individual heterogeneity--and potential conflict--among whatever language varieties have been acquired.' Since, he would argue, the foreign language is very much a carrier of values that do not always match those acquired in the native tongue, conflict is inherent in the language learning situation. One example, as Cazden notes, is the conflict between acceptability by the teacher and acceptability by peers in the classroom; we have a glimpse of this in excerpt 2 when the student switches to Swiss dialect as an act of solidarity with his peers and in rejection of the teacher's story about dictionary definitions. Excerpt 1 is another example of this 'conflit inhérent à la relation pédagogique' (Yves Chalon in Riley 1985:1), this time taking place within a given learner between his/her own narrative

voice and the teacher's abstract or 'alien' voice that he is forced to adopt. For Bakhtin, becoming communicatively competent means finding one's own voice and one's own role in the speech community of the classroom.

The evaluation. Evaluation in narratives occurs on two levels: that of the text (external evaluation) and that of its performance (internal evaluation). External evaluation occurs through corrections or acknowledgements by the teacher of the referential truth of the text ('He isn't waving, he is--the correct word is--he is wagging his tail'; 'yes he is happy' where 'yes' means: 'You're right, the dog is happy'). Internal evaluations enable the conversational partners to give each other feedback on their role in the conversation, i.e. how they present themselves through their story, in particular the degree of tellability of their responses. This evaluation can be negative: e.g. 'you said the reason why he is doing this' (meaning: 'but this is not what I want to talk about'); or positive: e.g. 'that isn't a relative clause...but that would be an answer, yes'. Evaluation can also bear on the appropriateness of the joint storytelling performance (e.g. teacher: 'who said the word cellar? that was your word wasn't it?' or student comment: 'awful').

The lesson is all the more successful the more students are given the opportunity to contribute to all three aspects of the narrative: its text, its performance, and its evaluation. Too often teachers and students in language lessons tend to neglect the performance aspect to the overwhelming benefit of the text, and evaluation consists almost exclusively of external assessments by the teacher of student text. For example, seen from a narrative perspective, excerpt 1 is neither 'right' nor 'wrong' teaching. It is only told in a manner that weakens its potential as a joint narrative, because it does not make use of the students' stories despite the teacher's invitation to join.

A narrative view of the language class. Described in narrative terms, a foreign language lesson oriented towards the development of communicative competence strives to achieve the richest possible performance of a given text, be it a grammar unit, the discussion of a reading or the transmission of cultural information. As one of the narrators, the teacher has both to select what story to tell and elicit, and decide how to orchestrate its telling. At the beginning levels, most of the teacher's energy is spent staging his/her students' voices and developing the speech community of the classroom through pair and group work, choral responses and all the resources of stress, intonation and nonverbal behavior. At later stages, the role of the teacher as narrator is mainly to bring to the fore and capitalize on the differences between the various classroom voices, fostering the growth of the students' 'capacity for individual response to language use' (Widdowson 1975:76).

A narrative perspective would entail focusing on the following important aspects of language teaching in classrooms.

(1) The classroom as both a window on the outside world and a 'self-referential, self-contained unit of communication, suspended from the immediate reality of social life' (Widdowson 1975), a speech community in its own right (Hymes 1973) with its own intertextual references.

(2) The lesson as 'representation'. Staging procedures--timing, spatial arrangement, prefacing activities, closures, wrap-ups--have to be viewed not

as a means by which the teacher manages students' learning, from behind the scene so to speak, but as a *mise-en-scène* that should be more and more explicitly evaluated and jointly decided upon by both teacher and students.

(3) The concept of tellability. Teacher and students have to constantly negotiate what is worth telling in class and what is not; how much use to make of anecdotes and digressions; how best to use the multiple and variable performance styles and modes of thinking of the participants in the narrative.

(4) Story-telling rights and duties. Joint performances have multiple voices with equal story-telling rights: it is up to each story teller to learn how to incorporate all of them into the story by finding potential relevancy in the contributions of others.

(5) Matters of truth. As we have seen, truth is not only in the propositional content of the text and its reference to an external foreign cultural world, it is also in the telling itself, in the performance. Personalization as a pedagogic goal does not mean that students should speak mainly about themselves and their daily lives; in a narrative sense, it means that they be allowed to have a voice as a 'persona' in the general narrative and that this persona be taken seriously. Students should be reminded that they need not answer in their name; 'grandmothers' need not be *their* grandmothers, 'future dreams' need not be *their* future dreams. Shared duplicity in fiction is part of constructing the common story. This keeps their self intact because it is not equated with their story-telling persona.

(6) 'Deep play'. Bruner (1976:57) calls 'deep play' an activity in which the personal stakes are high even though it is only a game. Foreign language lessons, whose personal stakes can be very high for the self-esteem of both teacher and students, can be viewed as a unique mix of deep seriousness and deep playfulness. Such a mix is found in conversational narratives, where conversational personae are brought together within a play frame in which they both recount external events and derive internal pleasure from the telling itself.

Conclusion. We started with a dichotomy of which many language teachers feel prisoners, between form and function, contextualized and decontextualized learning, real-life and academic necessities. I have suggested viewing the language lesson as conversational narrative, in which participants join in the construction of a common narrative in and through the foreign language. The richness and originality of that narrative depend on the opportunities that learners and teacher create to develop their own voice in its performance.

So the question is not what should we do on Monday morning but what kind of person will we be? Not what should teachers teach: form or function, grammar or free communication? But how can the foreign language story best be told? What will teacher and students allow themselves and each other to tell and how much variation in performance will they manage to elicit and integrate?

Hymes summed all this up very well ten years ago:

There is a current movement to go beyond collection and analysis of texts to observation and analysis of performance. That is essential, but perhaps

only the second moment of three... Continuous with the others, this third is the process in which performance and text live, the inner substance to which performance is the cambium, as it were, and crystallized text the bark. It is the grounding of performance and text in a narrative view of life. That is to say, a view of life as a potential source of narrative. Incidents, even apparently slight incidents, have pervasively the potentiality of an interest that is worth retelling... Not that the difference is in the topics. The difference is in the silences. There is a certain focusing, a certain weighting. A certain potentiality, of shared narrative form, on the one hand, of consequentiality, on the other (Cazden and Hymes 1978:32).

Note

I am grateful to Ruth Spycher for making available to me these excerpts which she recorded and transcribed.

Appendix

Excerpt 3 (From a 7th semester English class in Swiss high school (17-18-year-old students), recorded and transcribed by Ruth Spycher):

- T. any other ar any arguments for yes please
- S1. (er)
- S2. (em) many people go to a doctor even if they aren't seriously ill
- T. yes—is that an argument for private medicine?
- S1. (em) (coughs)
- T. so how do I phrase it on the blackboard, how do I put it (em): many people go to doctors if they aren't ill?
- S1. ja because they haven't to pay for it
- T. aha if you got to pay for it
- S1. (er)
- T. they must be ill—is that what you mean
- S1. no—but then you don't go to a doctor (mh) if it's only
- T. a minor illness aha—so—doctors (teachers writes on the BB while talking)
- S1. (heavy breathing)
- T. aren't—let's—can I rephrase it for you--overloaded with (S2 coughs) work because you only go to the doctor if you really need to—is that what you mean
- S1. no
- T. no? aha—then I'm afraid I don't understand (teacher laughs softly)
- S1. (em) when people don't have to pay for the doctor
- T. yea
- S1. (e) they go to see him even if they they have got a cold (pupil laughs softly after having said that)
- T. yes alright (teacher works on the blackboard) so (em) (a) ja (em) the doctor the doctor is—does not have to (teacher writes on the BB while talking) deal with patients who aren't seriously ill (S2 coughs) (e) is that what you mean Claudia?
- (gong)
- S1. (Claudia). yes (em) the state has to pay for these patients however
- T. yes—ok—that's an argument against public medicine—that's right—it's more an argument—even though it's true--yes ok ...

References

- Allwright, Richard L. 1980. Turns, topics and tasks: Patterns of participation in language learning and teaching. In: Discourse analysis in second language research, ed. Diane Larsen-Freeman. Rowley, Mass.: Newbury House. 165-87.

- Aston, Guy. 1986. Trouble-shooting in interaction with learners: The more the merrier? *Applied Linguistics* 7.128-43.
- Bateson, Gregory. 1976. A theory of play and fantasy. In: Bruner et al. (1976:119-29).
- Breen, Michael P., and C.N. Candlin. 1980. The essentials of a communicative curriculum in language teaching. *Applied Linguistics* 1.89-112.
- Bruner, Jerome S. 1976. Nature and uses of immaturity. In: Bruner et al. (1976:28-64).
- Bruner, Jerome S., A. Jolly, and K. Sylva, eds. 1976. *Play--Its role in development and evolution*. New York: Penguin.
- Cazden, Courtney. (forthcoming) Contributions of the Bakhtin circle to 'communicative competence'. Paper given at the First Conference of the American and British Associations of Applied Linguistics, University of Warwick, July, 1988. Forthcoming in *Applied Linguistics*.
- Cazden, Courtney, and D. Hymes. 1978. Narrative thinking and story-telling rights: A folklorist's clue to a critique of education. *Keystone Folklore* 22.1-2.
- Fröhlich, Maria, N. Spada, and P. Allen. 1985. Differences in the communicative orientation of L2 classrooms. *TESOL Quarterly* 19.27-57.
- Hymes, Dell. 1973. On the origins and foundations of inequality among speakers. *Daedalus*, 59-85.
- Kramsch, Claire. 1984. *Interaction et discours en classe de langue*. Paris: Hatier-Credif.
- Kramsch, Claire. 1988. To group or not to group? Le parler, le faire et le dire en travail de groupes. *Les Langues Modernes* 6.16-23.
- Long, Michael H., and C.J. Sato. 1983. Classroom foreigner talk discourse: Forms and functions of teachers' questions. In: *Classroom-oriented research in second language acquisition*, ed. Herbert W. Seliger and M.H. Long. Rowley, Mass.: Newbury House. 268-86.
- McNeill, David. 1985. Language viewed as action. In: *Culture, communication and cognition*, ed. James V. Wertsch. Cambridge: Cambridge University Press. 258-70.
- Nold, Guenther. 1983. Grundzüge eines lernfördernden Sprachverhaltens des Lehrers im Fremdsprachenunterricht. In: *Kommunikation im Klassenzimmer und Fremdsprachenlernen. Protokoll eines Werkstattgesprächs des Goethe House New York im September 1982*, ed. Manfred Heid. Munich: Goethe Institute. 291-318.
- Polanyi, Lydia. 1979. So, what's the point? *Semiotica* 25.208-41.
- Riley, Philip, ed. 1985. *Discourse and learning*. New York: Longman.
- Rivers, Wilga M. 1975. The natural and the normal in language learning. *Language Learning*. 1-8.
- Rivers, Wilga M., ed. 1987. *Interactive language teaching*. Cambridge: Cambridge University Press.
- Soulé-Susbielles, Nicole. 1988. Mais que peuvent-ils donc se dire? *Explorations dans le travail de 'paires'*. *Les Langues Modernes* 6.24-31.
- Tannen, Deborah. 1984. The oral/literate continuum in discourse. In: *Spoken and written language. Exploring orality and literacy*, ed. Deborah Tannen. Vol. IV in the series *Advances in discourse processes*, ed. Roy O. Freedle. Norwood, N.J.: Ablex.
- Swain, Merrill. 1985. Communicative competence: Some roles of comprehensible input and comprehensible output in its development. *Input in second language acquisition*, ed. Susan M. Gass and C. G. Madden. Rowley, Mass.: Newbury House. 235-53.
- Watson, Rita, and D.R. Olson. 1987. From meaning to definition: A literate bias on the structure of word meaning. In: *Comprehending oral and written language*, ed. Rosalind Horowitz and S. Jay Samuels. New York: Academic Press. 329-54.
- Widdowson, Henry. 1975. *Stylistics and the teaching of literature*. London: Longman.
- Young, Katherine. 1982. Edgework: Frame and boundary in the phenomenology of narrative communication. *Semiotica* 41. 277-315.

Language teaching and theories of language

Charles A. Ferguson
Stanford University

Abstract. The relation between, on the one hand, linguistic theory and, on the other hand, the study of second language acquisition (SLA) and language teaching has been regarded in different ways. Some linguists have viewed linguistic theory as a primary theoretical base to be applied in language teaching and have viewed some of the phenomena of SLA as a validation of linguistic theory. In the early structuralist period they regarded Contrastive Analysis as the basis of SLA research, and in the current phase of the generative period they regard a parameter-setting model of Universal Grammar as the basis of SLA research. The viewpoint presented in this paper is a restatement of the author's view that SLA research and aspects of language teaching are a primary source of data and generalizations that can test and modify linguistic theories, and thus, rather than being a field for confirmation and application they may offer an important contribution to linguistic theory construction.

The relation between teaching foreign languages and the construction of linguistic theory can be viewed in various ways. The two activities have different goals and proceed in different ways, but they overlap, for example, in the notion of 'a language' as a total system of some sort, a set of knowledges and skills shared by a social group and acquired by the individuals in it. This notion is, however, problematic. For example, there is no consensus on how to distinguish between pairs of language varieties that belong to one language and pairs that belong to two different languages. Also, the relation between individual and social language repertoires is unclear, and the way to incorporate the 'ordered heterogeneity' of the variationists into a total system is still undiscovered. But if we put these problems aside, and assume there is a natural entity that corresponds in a rough way with what both linguistic theorists and language teachers regard as 'a language,' the relation between the two fields can still be regarded in various ways. One of the great merits of the Georgetown University Round Tables on Languages and Linguistics has been the inclusion of different perspectives on this issue. In this paper I want to present two views of the relationship. One of them has been fairly popular among linguists but also among some language teachers and researchers in second language acquisition (SLA). The other is less often expressed, but it is a view that I hold, and having stated it at earlier Round Tables (e.g. 1963, 1968), as well as elsewhere, I welcome the chance to offer it again for consideration.

The first view is that the findings of linguistic research, both the description of particular language varieties and--most especially--linguistic theory itself, have much to offer the language teacher, and that the experience of language teaching and research in SLA have little or nothing to offer linguistics. The second view is that the two fields have much to offer each other and, in the present state of their respective developments, the findings

of SLA research, including the observed phenomena of language teaching, constitute one of the most important sources for the testing of linguistic theories and indeed for the construction and elaboration of linguistic theory. Both views, as well as some others, are compatible with the concept of 'applied linguistics' in its usual sense of the 'application of the methods and results of linguistic science to the solution of practical language problems' (Ferguson and Morgan 1959). The solution of practical language problems, such as the creation of new writing systems, translation from one language to another, or teaching a foreign language, always requires knowledge drawn from fields other than linguistics, but the nature of human language as the linguist understands it seems always to be involved at some level and often crucially so.

The view that the flow of knowledge is unidirectional--linguistic theory to language teaching--is perhaps a natural derivative of linguists' perpetual excitement about theoretical insights on the nature of human language and their typical pride in the advanced state of linguistic theorizing. The great example of this view in the 1950s and 1960s was many linguists' conviction of the validity of contrastive analysis (CA) theory. The great example in the 1980s is many linguists' conviction of the validity of a parameter-setting universal grammar (UG) theory.

CA theory. The CA theory claimed that languages of the world differ in an indefinitely large number of ways and that each language has its unique characteristic structure in terms of which the speakers operate. Further, speakers of a language tend to hear another language and attempt to produce utterances in it in terms of the structure of their own language, thus accounting for their 'accent' in the L2. (I am using 'accent' here to refer to any phenomena that mark the L1 speakers of an L2, as when a German learner of English pronounces [s] for /θ/, says *I study English since five years* for *I've been studying English for five years* or uses *eventually* to mean 'in case it should happen' instead of 'ultimately'.) Finally, the theory claimed that a contrastive analysis of the L1 and L2 will tend to indicate the 'trouble spots' for the learner, the points of special difficulty and the types of errors likely to be found at them. That was a promising theoretical position to take, for which anecdotal evidence was extensively available. As I have pointed out elsewhere (Ferguson 1985), the claims of this theoretical position were essentially linguistic in nature. They were basically claims about the nature of human language in general and linguistic aspects of human language processing. CA also, of course, claimed to make predictions of a sort about the behavior of L2 learners.

To a detached observer of the linguistic scene, proposing the CA theory would seem to have offered an ideal kind of theoretical claim, namely one that could readily be tested and modified in response to the results of the testing. Our mythical detached observer could have reasonably assumed that linguistic theorists would welcome the chance to make predictions of L2 errors on the basis of contrastive structural mismatches and, by careful examination of the errors, to sharpen the descriptive categories and analytic tools of their theory.

We all know this is not what happened. CA was not seen as linguistic theory to be tested, and L2 error data were not seen as a valuable source of

evidence for such testing and certainly not as a prime contribution to the construction of new and improved linguistic theory. Instead, CA was taken as established linguistic theory that could be applied to such practical language problems as foreign language teaching. At best, linguists thought that the data from SLA research might offer gratifying confirmation of the theory.

Instead of developing an extensive line of research using the error data to build phonological and syntactic theory, researchers regarded CA as a tool for pedagogical analysis. They produced innumerable contrastive analyses, a much smaller number of systematic error analyses in relation to CA predictions, and almost no feedback from SLA research into linguistic theory itself. CA theory had been made explicit by Prague School structuralists in the 1930s (e.g. Polivanov 1931) and was widely held by American structuralists in the 1940s, but it was not until 1966 that the first report of testing CA theory appeared in a major American linguistic journal (Brière 1966).

Even when problems of error interpretation arose that raised immediate, direct questions of linguistic theory, they were usually not explored. One example will suffice. In Lado's book *Linguistics across Cultures* (Lado 1957), which was the best known exposition of CA for more than a decade (until Di Pietro 1971 and James 1980), the author noted the problem of different substitutions for English /θ/ by speakers of languages that lacked an equivalent phoneme. Speakers of some languages tended to produce [s], speakers of other languages produced [t], and speakers of still other languages fluctuated, even though all the languages had both an /s/ and a /t/ of some kind in their phoneme inventories. The L1 languages mentioned by Lado were Japanese, Spanish, Tagalog, and Thai; English was the L2 in each case. Although many factors may be responsible for errors of this kind, it is certainly plausible that the phonological structures of the respective L1 and L2 are critical. The then current phonological theory offered no help. Lado acknowledged the inadequacy of CA in this case, but did not suggest phonological research.

Sporadic attempts to deal with this issue for particular pairs of languages appeared in structuralist literature in Europe and America, but no one saw the question as a stimulus to systematic general theory-oriented research. Ritchie (1968) examined this question of differing substitutions for English /θ/ by Russian speakers and Japanese speakers and suggested an explanation for the particular difference pair in terms of the Chomsky-Halle generative phonology of the time, but he made no attempt at more general principles. Schmidt (1977) offered a sociolinguistic explanation for the fluctuation in Arabic substitutions for English /θ/. Such isolated studies still continue, as Cairns (1988) investigates the mismatch in Japanese and English /s/s that is held to account in part for Japanese substitutions for English /θ/. But no one has posed this question as a serious research task: to find theoretical principles accounting for cross-language differences in phonological errors for the same target sound from apparently the same possible substitutions. If phonologists had seen CA as a linguistic hypothesis rather than a pedagogical solution, it seems likely that some intensive research on an array of interesting cases might well have led to 'richer and more realistic forms of phonological analysis' (Ferguson 1985). No wonder Bolinger commented in his introduction to Di Pietro's book (Di Pietro 1971:viii) 'what we know about interference is

empirical knowledge and does not depend on theories, which themselves fail precisely to the extent that they are unable to account for interference.'

The CA theory was inadequate in many respects, both as linguistic theory and as a theory of SLA. It made no explicit allowance for universal properties of human language. It omitted consideration of parallels in first language acquisition. It ignored nonlinguistic explanatory factors. It made no explicit provision for intermediate steps in the development toward target structures.

The CA theory, however, had the great merit of highlighting the phenomena of linguistic transfer or interference. If these phenomena had been investigated more energetically by linguistically oriented researchers as an area of direct relevance for understanding the nature of human language, it would not have taken so long to reach some of the tentative conclusions we now have. It was not until 1983 that a full book of readings on the topic of linguistic transfer finally appeared (Gass and Selinker 1983), based on a conference devoted to the topic. Their book assumes that 'language transfer is an important aspect of the second language acquisition process and investigates what the constraints are on its occurrence' (7-8).

Research findings suggest that transfer is more likely in phonology than in syntax (Ioup 1984), more likely in the semantics of grammatical categories than in the forms themselves, more likely in formal classroom instruction than in untutored, 'natural' acquisition, more likely among older learners than younger ones, more likely in situations where the L1 continues in use alongside the L2 than in situations where the use of the L2 is dominant (Sridhar 1989). To my knowledge no theory of language, linguistic or otherwise, provides an explanation for these tentative findings. It is tempting to agree with Myhill's conclusion in a paper on syntactic transfer in SLA (Myhill 1982): 'At the moment, interference is where we find it, and as research continues we will have more of an idea where to find it.' The experiences of language teachers and learners constitute a major source of evidence for confirming, modifying, or disconfirming theories of language, and a major source of stimulus and challenge for linguistic theory builders.

UG theory. The current example of the view that the flow of knowledge is from linguistic theory to SLA and language teaching is what I will call the UG theory. (I am using this name to refer to the government and binding model of principles and parameters, with some apology to other theories that also claim the concept of universal grammar.) The UG theory assumes that there are principles underlying all natural languages that determine the basic grammar of any language. Included among these principles are parameters that are set by experience. Setting one value of a particular parameter yields one language and a different value of the same parameter yields a different language. The whole system has been summarized in metaphorical terms (Chomsky 1986 as cited in Flynn 1987:29): 'We may think of UG as an intricate system associated with a finite set of switches, each of which has a finite number of positions. The role of input is relegated to that of setting the switches. When they are set, the system functions.' In the process of second language acquisition, the learner has already acquired (or in the case of younger learners is in the process of acquiring) the parameter values of the L1, and must acquire the values of the L2. In the case of a match between

the two sets of values the acquisition is facilitated. In the case of a mismatch the acquisition of the L2 is to some extent disrupted. In principle this theory, like the CA theory, also claims that a comparison of L1 and L2 with respect to one or more parameters will tend to indicate the 'trouble spots' for the learner and the types of errors likely to be found at them.

This is a promising theoretical position to take. Unlike some other linguistic theories we have had in the meantime, it provides for a relationship between linguistic theory and second language acquisition. As in the case of CA theory, to a detached observer of the linguistic scene, proposing this UG theory would seem to offer an ideal kind of theoretical claim. Our mythical detached observer could reasonably assume that linguistic theorists would welcome the chance to make predictions of L2 errors on the basis of differential settings of various parameters, and by careful examination of the errors to sharpen the formal categories and improve the specification of various parameters.

It seems very likely, however, that as in the case of CA theory, the linguists, SLA researchers, and some linguistically oriented language teachers will view this theory not as a linguistic theory to be tested by SLA but as a linguistic theory to be applied as a paradigm for SLA research which would, incidentally, serve to confirm the theory.

In their generally excellent account of the recent history of relations between linguistic theory and SLA research, Newmeyer and Weinberger (1988) make the observation: 'If such notions as "markedness," "the sonority hierarchy," "parameterized grammars," and so on can be shown to be necessitated by the facts of second language acquisition, then their incorporation into UG is ipso facto supported.'

As far as it goes, this view is hard to disagree with. What one would welcome, however, would be the acknowledgement that if SLA research gives counterevidence for such notions, then the current UG theory should be carefully reexamined. Even more welcome would be the suggestion that SLA research might well lead to 'principles' or 'parameters' not yet hypothesized in UG theory but worth considering in linguistic theory construction. It would, I suppose, be too much to hope for that SLA research might lead to a substantial revision and improvement of the theory itself.

Let me take a single example of a problem of error interpretation that raises immediate, direct questions about linguistic theory. All varieties of Arabic, as far as I know, have obligatory resumptive pronouns in relative clauses, and all varieties of Arabic, as far as I know, require a relative connector when the antecedent is definite and do not permit the relative connector when the antecedent is indefinite. Arabic says, 'The man that I saw him yesterday visited me today' and 'A man I saw him yesterday visited me today.' Anyone who teaches English to speakers of Arabic will have noticed that the Arab learners tend to insert resumptive pronouns in English relative clauses where they don't belong, especially in the earlier stages of acquisition but often persisting a long time. Similarly, the speaker of English who learns Arabic often fails to insert resumptive pronouns where they belong, especially in the early stages of acquisition but often persisting a long time (cf. Schachter 1974, Eckman 1977). The English-speaking learner of Arabic occasionally inserts a relative connector with an indefinite antecedent, but on the whole the

consequences of the mismatches in expression of the relative connector are not readily detected either way because of the variable presence of the *that* in English relative clauses.

Although many factors may be responsible for L2 errors of this kind, it is certainly plausible that the syntactic parameter settings of the respective L1 and L2 are critical. Current UG theory offers no immediate help. It is not even clear what parameter or parameters are involved and what the respective settings might be. Many languages have resumptive pronouns and many don't, and the details of obligatory and optional resumptives vary from one language to another (at least partly in relation to the accessibility hierarchy). This is a typical example of a question that can be proposed as a serious research task: to find theoretical principles accounting for a fairly clearcut syntactic transfer phenomenon. Hyltenstam's fine study of learners' use of resumptive pronouns in L2 Swedish is a good start (Hyltenstam 1984), although it makes no mention of UG or parameters.

From my perspective the UG theory has two great merits. First, it highlights systematic differences across languages that yield linguistically significant language typologies, and, second, it claims a systematic relationship between language structure and second language acquisition. Also, it clearly improves on the CA model in that it rests firmly on the notion of universal properties of all languages.

Of course, like any linguistic theory it also has weaknesses. An obvious one is its failure to acknowledge the enormous amount of language-particular phenomena. The acquisition of the whole lexicon, for example, with all its intricate constraints and idiosyncrasies, many details of morphophonemics, and so-called 'crazy rules' in phonology remain outside the model. Any theory that accounts for the successful first language acquisition of the child by virtue of universal grammar must also account for the similarly successful acquisition of these language-particular phenomena. Some UG proponents have made such astonishing claims as that 'the virtually unerring rapid learning of lexical material' is to be accounted for by the fact that 'the concepts for which words and morphemes are to be learned are themselves universal (for the most part)' (Hale 1988:30). Anyone who has taught a foreign language needs a better account of vocabulary learning than that as well as an explanation for L1-L2 transfer in the mismatches of lexical semantics.

UG theory shares with CA theory the failure to predict in any way the path of learning from early errors to later close approximations to the target structures: the factors involved cannot all be nonlinguistic. But the most serious problem with UG is the vagueness and indeterminacy in characterizing the notion of parameter. Just what constitutes a parameter? Are there subparameters? Clusters of parameters? Systemic consequences throughout the grammar of particular parameter settings? Interactions among various parameters? The pro-drop or null subject parameter is the one that has probably had the most extensive exploration, but we are nowhere near a specification of exactly what that parameter is and what settings are possible. Certainly null subject languages differ widely in relevant syntactic detail and it seems doubtful that there is a simple linear scale of settings. It is here that SLA research, including the experience of language teachers, offers some of the most direct insights into any study of parameter setting. Incidentally, SLA

error analysis may offer suggestions for language teachers even when a sound theoretical basis is lacking.

I suppose the main point of this paper is that the flow of knowledge between linguistic theory on the one hand and SLA research and language teaching on the other is bidirectional. It goes both ways, and in the present state of development of the UG theory it seems possible that the SLA researchers and linguistically oriented language teachers may have more to offer the linguistic theorists than vice versa. The main fear underlying this paper is that the linguistics and language teaching communities may make the same mistake with UG that was made with CA, overvaluing the linguistic theory, undervaluing the SLA research, and failing to have the two fields feed each other.

In the case of CA theory, many researchers were led to reject the whole theory because it was not seen as linguistic theory to be tested but as a finding or 'truth' of linguistics to be applied. The important insight of the close relation between typological differences in languages and the phenomena of linguistic transfer had to be painfully reintroduced into linguistic theory decades later when it could have come about by using feedback from SLA research.

In closing, I cannot maintain that this view of the interdependence of linguistic theory and SLA research is mine alone. Something similar is found in the Bolinger introduction I cited earlier (Bolinger 1971), in Comrie (1984), and in the Lightbown paper on linguistic theory and SLA in the 1986 Georgetown University Round Table (Lightbown 1986). It appears--though somewhat hidden--in contributions to the Flynn and O'Neil volume *Linguistic Theory in Second Language Acquisition* (Flynn and O'Neil 1988), and it is to be found also in the volume that Huebner and I have edited, *Crosscurrents in Second Language Acquisition and Linguistic Theories* (Ferguson and Huebner forthcoming). The hope underlying the present paper is that more people will come to share this view.

References

- Bolinger, D. 1971. Introduction. In: *Language structures in contrast*, by R. Di Pietro. Rowley, Mass.: Newbury House.
- Brière, E. 1966. An investigation of phonological interference. *Lg.* 44.768-96.
- Cairns, R. S. 1988. Hearing and articulating /s/ in English and Japanese. *Cross Currents* 15.47-56.
- Chomsky, N. 1986. *Knowledge of language: Its nature, origin, and use*. New York: Praeger.
- Comrie, B. 1984. Why linguists need language acquirers. In: *Language universals and second language acquisition*, ed. W. E. Rutherford. Amsterdam: John Benjamins.
- Di Pietro, R. 1971. *Language structures in contrast*. Rowley, Mass.: Newbury House.
- Eckman, F. R. 1977. Markedness and the contrastive analysis hypothesis. *Language Learning* 27.315-30.
- Ferguson, C. A. 1963. Linguistic theory and language learning. In: *Georgetown University Round Table on Languages and Linguistics 1963*. Washington, D.C.: Georgetown University Press. 115-22.
- Ferguson, C. A. 1968. Contrastive analysis and language development. In: *Georgetown University Round Table on Languages and Linguistics 1968*. Washington, D.C.: Georgetown University Press. 101-12.
- Ferguson, C. A. 1985. Contrastive analysis: A linguistic hypothesis. In: *Scientific and humanistic dimensions of language: Festschrift for Robert Lado*, ed. K. R. Jankowsky. Amsterdam: John Benjamins.

88 / Georgetown University Round Table on Languages and Linguistics 1989

- Ferguson, C. A., and T. Huebner, eds. (forthcoming) *Crosscurrents in second language acquisition and linguistic theories*. Amsterdam: John Benjamins.
- Ferguson, C. A., and R. Morgan, Jr. 1959. *Selected readings in applied linguistics*. (Linguistic Reporter Supplement 2) Washington, D.C.: Center for Applied Linguistics.
- Flynn, S. 1987. Contrast and construction in a parameter-setting model of L2 acquisition. *Language Learning* 37.19-59.
- Flynn, S., and W. O'Neil, eds. 1988. *Linguistic theory in second language acquisition*. Dordrecht: Kluwer.
- Gass, S., and L. Selinker eds. 1983. *Language transfer in language learning*. Rowley, Mass.: Newbury House.
- Hale, K. 1988. Linguistic theory: Generative grammar. In: Flynn and O'Neil (1988).
- Hyltenstam, K. 1984. The use of typological markedness conditions as predictors in second language acquisition: The case of pronominal copies in relative clauses. In: *Second languages: A cross-linguistic perspective*, ed. R. W. Andersen. Rowley, Mass.: Newbury House.
- Ioup, G. Is there a structural foreign accent? A comparison of syntactic and phonological errors in second language acquisition. *Language Learning* 34.1-17.
- James, C. 1981. Contrastive analysis. Harlow, Essex: Longman.
- Lado, R. 1985. *Linguistics across cultures*. Ann Arbor: University of Michigan Press.
- Lightbrown, P. 1987. The influence of linguistic theory on language acquisition research: Now you see it now you don't. In: *Georgetown University Round Table on Languages and Linguistics 1987*. Washington, D.C.: Georgetown University Press. 130-42.
- Newmeyer, F., and G. Weinberger. 1988. The ontogenesis of the field of second language research. In: Flynn and O'Neil (1988).
- Polivanov, E. D. 1931. La perception des sons d'une langue étrangère. *Travaux du Cercle Linguistique de Prague* 4.79-96.
- Ritchie, W. C. 1968. On the explanation of phonic interference. *Language Learning* 18.183-97.
- Schmidt, R. W. 1977. Sociolinguistic variation and language transfer in phonology. *Working Papers in Bilingualism* 12.29-95.
- Sridhar, S. N. 1989. The empirical study of a non-native variety of English. Paper presented at the Presession on English in Bilingual Speech Communities, Georgetown University Round Table on Languages and Linguistics 1989. Washington, D.C.

Cohesion and coherence in the presentation of machine translation products

Muriel Vasconcellos

*Pan American Health Organization
Washington, D.C.*

Abstract. Different degrees of human intervention can be applied to the preparation of machine translation (MT) products in readying them for their ultimate use.

It is suggested here that differences in levels of postediting are associated to some extent with cohesion and coherence. The posteditor is essentially an interpreter of discourse: much of the postediting task involves either employing devices to ensure that the surface pieces of the discourse are connected in meaningful ways or else adjusting the reading of each item against the interpretation of others until the entire underlying text is made to cohere.

The presence of cohesion and coherence is examined in three versions of the same machine translation: the raw output, a lightly postedited version, and the final, fully postedited product.

1 Perspective on MT postediting. Machine translation can be delivered to the end user as raw output, of course, and it can also be postedited to varying degrees. There is much to be learned from a look at the doctoring that is done when the machine stops working and the human user takes over. Certainly a linguistic study of intervention in the MT product at different levels of refinement can help us to prioritize our strategies. By stratifying the types of corrections that are made, we can begin to orient postediting policy so that today's MT systems are used more effectively, and we can also contribute to the improved performance of the systems of tomorrow.¹

In practice, time and cost constraints often lead to situations in which postediting is curtailed to one degree or another. Depending on the purpose of the translation, nuancing may be traded off for expediency and economy. The most drastic curtailment, of course, is no postediting at all, as is sometimes the policy with translations for information only. Usually, however, even with informative translations there is some type of human intervention. Newman (1988), based on experiments with the SYSTRAN and LOGOS MT systems, has recommended limiting information-only postediting to the replacement of foreign words--words not found in the MT dictionary. Somewhat more intervention is practiced at the U.S. Air Force Foreign Technology Division, where 'partial postediting' addresses seven types of target errors (Bostad 1987). SYSTRAN's Russian-English translations are passed through an automatic postprocessor (EDITSYS) which produces warning flags; whenever any of the seven types of error occurs in the output, the human operator is alerted by a flashing line across the screen. The corrections elicited by EDITSYS will affect, on average, about 20% of the output (Bostad 1987:438). As with most 'information-only' translation, the material handled by the Air Force covers a broad range of subjects and comes

from a wide range of sources. This is the opposite of constrained input in a highly limited domain, where MT systems may be able to handle most problems at the level of the algorithm and generate a usable translation that requires very little correction. This latter situation is exemplified by METEO 2, the system that translates Canadian weather forecasts; as of mid-1988, interventions were down to the point that only 3.4% of the text was being affected (Chandioux p.c.).

None of these applications, however, is what you might call 'mainstream' translation. In the everyday world, by far the greatest demand is for translations of general and technical material that leave no doubt as to the meanings intended by the original author. To produce such translations, given the current state of MT art, may require a somewhat more intensive human review than what has just been described. Still, it would be useful to distinguish levels of 'light' and 'full' postediting. At the European Commission in Luxembourg, for example, 'rapid' postediting of SYSTRAN machine translations was sanctioned under a project launched in May 1982 (Wagner 1983). Differences between rapid and conventional postediting were later examined by Löffler-Laurian (1986).

For our work at the Pan American Health Organization (PAHO), where we have been postediting machine translation for nearly ten years, it would be of great practical value to identify the differences between two such levels and to systematize their implementation. Moreover, to the extent that we are able to relate the distinctions to broad linguistic principles, our findings may be of more general interest for MT development and for translation theory as well.

2 Approach to a definition of levels. My hypothesis is that the differences in translation 'quality' at the levels of rapid and full postediting can be correlated, at least partially, with syntactic corrections and degrees of COHESION, on the one hand, and COHERENCE on the other.

Of course, raw MT output, as well as that which is checked only minimally for predictable trouble spots, is bound to have some syntactic problems that need to be corrected. Once these have been dealt with, it is likely that the product may still be further improved: cohesive devices can be introduced that will establish clearer connections between the pieces, and more refined interpretations can be made of the nuances that help to convey the author's intentions to the target audience.

For the rapid postedit, Löffler-Laurian (1983) proposes that revision should concentrate on 'vocabulary changes', especially in domains for which the dictionary has not been highly developed: translations should be supplied for not-found words, and erroneous glosses should be corrected. Also, passages that are incomprehensible should be repaired. These are useful criteria. In addition, on the basis of our experience at PAHO, I would say that at this level many devices can be invoked which will enhance the cohesiveness of the text.

The full postedit, in turn, involves modifications that will bring out nuances and enable the reader to grasp the complete significance of the text. It makes the difference between a translation that is merely passable and one that is appropriate for the most demanding of circumstances. Löffler-Laurian (1983) has offered a set of four guidelines and twelve specific rules for the

posteditor working at this level.² Here the PAHO experience points, further, to interpretations leading to improved coherence.

In testing the hypothesis about the respective roles of cohesion and coherence, it is important that we work from a clear definition of each of these terms.

3 Concepts. Widdowson (1979:87) defines cohesion as 'the overt structural link between sentences as formal items', and coherence as 'the link between the communicative acts that sentences are used to perform'. He goes on to suggest that cohesion is the propositional relation between the parts of a discourse, whereas coherence is the illocutionary relation. For present purposes, the definitions of both concepts have been broadened to apply to relations *within* sentences or communicative acts as well as between them. Thus, cohesion is taken to refer to ties between elements manifest in the surface structure of the discourse, while coherence has to do with the interpretation of connectedness in the underlying text.³

3.1 Cohesion. Cohesion is easier to describe than coherence, and easier to recognize. The devices can be specified, and when applied to translation they can yield considerable payoff in terms of understandability.

According to Halliday and Hasan (1976:4), cohesion occurs when an element in discourse cannot be effectively decoded without invoking another element in the text or the discourse situation. 'When this happens, a relation of cohesion is set up, and the two elements, the presupposing and the presupposed, are thereby at least potentially integrated into a text'. They identify the following types: reference, substitution/ellipsis, lexical cohesion, and conjunction. Since these headings will be applied below to some of the corrections that are made in MT postediting, they are elaborated here in some detail.

In the case of reference, an element in discourse relies on some other element for its interpretation: information must be recovered about it-- either a referential meaning or the identity of a particular thing--in order for it to be decoded. Personal and possessive pronouns, for example, set up pronominal reference. Demonstrative reference is established by demonstrative pronouns and also by the definite article *the*. Comparative reference involves identity, similarity, difference, or quantitative or qualitative relations between discourse entries. The referent may be present in the discourse situation rather than the text, in which case the reference is exophoric. Cohesion is created by the fact that the same concept enters the discourse a second time, and the cohesive tie is the connection between the two occurrences (31).

Substitution is 'the replacement of one item by another' (88). The second item, or substitute, establishes a cohesive link with the first. Whereas reference is a relationship between meanings, substitution is a relationship between linguistic items. The substitute is used to avoid repetition. In English, NPs can be replaced by *one(s)* or *same*; verbs by *do* (+*so/it/that/the same/likeywise*), *be* (+*so/it/that*), *have* (+*to*); and clauses by *so* or *not*. Unlike reference, substitution cannot be exophoric; it can only involve the elements expressed in the discourse proper.

Ellipsis (142) may be seen as a special subtype of substitution in which a linguistic item is replaced by nothing. There is cohesion with the zero element in the same way as there is with the substitute.

Lexical cohesion (274) connects discourse entries not through grammatical resources, as above, but rather through lexical choice. A synonym, a broader or narrower term, or a related term revives a concept in the discourse. There is also cohesion between any pair of lexical items that belong to the same ordered set or paradigm (*Tuesday...Thursday, north...south*) and between items that often cooccur--for example, *blade...sharp, garden...dig, try...succeed, king...crown, boat...row* (285).

Conjunction is different from the other four types of cohesion because attention is focused on the meaning of the cohesive relation itself rather than on the elements that are tied together (226-227). There is a large inventory of cohesive relations under the broad heading of conjunction, and the authors have classified them according to their function: additive ('and', 'or else', 'furthermore', 'for instance', 'similarly', 'on the other hand'), adversative ('but', 'nevertheless', 'in fact', 'on the other hand', 'instead', 'rather', 'in any case'), causal ('therefore', 'with this in mind', 'it follows', 'in that case', 'otherwise'), and temporal ('next', 'at once', 'meanwhile', 'finally', 'up to now', 'in short').⁴ Presumably the authors' lists could be expanded to include such discourse markers as 'oh', 'well', 'y'know', 'I mean' (Schiffrin 1987).

In addition to using the foregoing devices, which can be formulated quite explicitly on the basis of grammatical and lexical properties, cohesion involves developing the overall fabric of the text through the distribution of new and old information and through the staging effect created by message themes (Halliday and Hasan 325, Halliday 1967-68, Vasconcellos 1985, 1986a, 1986b).

3.2 Coherence. Unlike cohesion, coherence underlies the discourse and has no predictable reflex in surface structure. Whereas cohesion has to do with relations between surface linguistic forms and between propositions, coherence involves connectedness within the communication act itself. The speaker/writer is now seen as communicator, and the listener/reader as interpreter.

The progress of a discourse is determined by the communicator's choices of meanings to be focused on. In turn, the interpreter of a discourse (in our case the posteditor) must be able to decide for each entry in the discourse which meaning type,⁵ and within it which specific meaning among possible alternatives, is intended. If the posteditor's interpretation matches the author's intention, the translation is fully successful--although in reality this success is apt to be achieved only to an approximate degree.

Communicators and interpreters assume that a text is coherent. Coherence is observed, and therefore defined, more through its absence than its presence. Lack of coherence may be illustrated by the following example (van Dijk 1972):

- (1) We will have guests for lunch. Calderón was a great Spanish writer.

Despite the strong tendency to assume coherence, it is difficult for an interpreter to see any connectedness between the two entries. Coherence is present, on the other hand, in a similar sequence:

- (2) You ought to read *Wombats Galore*. Bruce McQuarrie is a great author.

even though in fact it is nonsense contrived precisely to make this point (Stubbs 1983:124).

Some authors would assign part of semantic connectedness to cohesion,⁶ but the position taken here is that coherence, rather than cohesion, underlies the interpretations of textual meaning. For Sanders (1987), it is coherence which provides the communicator with the cognitive basis for formulating discourse entries so that control is exercised over the way he or she is understood (7). From the perspective of the interpreter, who in the case of written text is distanced from the author at least in time if not in space, readings have to be adjusted back and forth as the discourse unfolds until each discourse entry has a specific interpretation that fits with what went before (84).

A single word may constitute a discourse entry, and its reading has to be adjusted against others in the context until they are made to cohere as much as possible. Sanders illustrates this process with a well-known sentence:

- (3) Colorless green ideas sleep furiously.

At first we are struck by blatant incoherence. This is because we give an unmarked interpretation to each of the terms. However, if we force ourselves to assume that the message is coherent, we can try to read different meanings into the components until a coherent interpretation of the whole is arrived at. Each term is examined for its range of possible meanings, and the various options are tested against the surrounding context. To start with, by looking ahead we know that *colorless* does not readily apply to the upcoming concepts of *green* and *ideas*, so we backtrack and interpret it as 'lackluster'. We then rethink the meaning of *green* and reject the more usual one of 'a color' in favor of 'unripened'. And so on. Sanders' result is:

- (3') Lackluster unripened ideas lying dormant are volatile.

The process that Sanders describes is constantly exercised in translation, especially in the postediting of machine output. Postediting is an ongoing process of interpretation, since the pieces of the target language are already given. The job of the posteditor is to examine these pieces, make a 'specific interpretation' of the meaning intended by the author, and adjust the wording so that the text becomes more coherent. The computer can and often does generate a set of pieces which an interpreter can appreciate as a fully understandable translation which is both grammatically and discoursally well-formed.⁷ But this judgment has to be made by the posteditor, and such sentences may be intermixed with others that are less felicitous.

4 Cohesion and coherence in postediting. In our work at the Pan American Health Organization with Spanish-English and English-Spanish MT (SPANAM™ and ENGSPAN™, respectively⁸), we have found that a conventional postedit is usually accomplished in two passes: a first 'rough draft' and then a final polishing. This has been our typical experience over the last ten years. We would like to know now whether these two passes could be made to correspond to linguistically describable levels.

What is really done in each of the two passes? Are there linguistic criteria that distinguish one level from the other? And finally, are there purposes for which raw MT and first-pass MT are adequate? Up to now the answers to these questions have eluded us. No one has been able to tell posteditors exactly what to look for. Of course, with many of the changes that are introduced there is general agreement on the need for something to be done--if not on the solution. But with other changes there is debate about whether they are essential or even worthwhile at all. So far, it has been difficult to clarify how they contribute to making the translation more explicit--and therefore more useful.

To address these questions, I singled out one of the jobs in our regular Spanish-English production stream for which it was possible to reconstruct three different versions: the raw output (presented in side-by-side form as Appendix A and in target-only form as Appendix B), a first-pass postedit (Appendix C), and the final translation (Appendix D). The complete text was a 7,000-word report on the status of nutrition in Latin America, a subject on which SPANAM has often been exercised in the past. The first 312 words, which are fairly typical of the rest of the document, were examined in depth and are discussed at each of the two levels in the following sections.⁹

4.1 The first pass. For both the first pass and the final translation, the changes that had been made were grouped under three broad headings: syntactic corrections, cohesive devices, and interpretations for coherence. An effort was made to assign all the changes, including lexical choices, to one or another of these categories.

At the level of the first pass, the posteditor made a total of 33 changes (shown in Appendix B).^{10,11} The distribution was as follows:

Syntactic corrections	9	27%
Cohesive devices	21	64%
Interpretations for coherence	3	9%
		100%

Syntactic corrections. Three of the syntactic changes were merely punctuation: one a comma to mark a nonrestrictive relative clause [line 13], another a comma to match an existing comma for a parenthetical phrase [line 17], and the last a hyphen [line 22]. Two others also involved the further marking of a nonrestrictive relative clause: omission of *and* as a translation of the Spanish clause-marker *y* and substitution of *which* for *that* [both on line 13]. Two were corrections in prepositional government (*place demands on* [line 13] and *suffer from* [line 21]). An adjustment was made to accommodate the fact that *contribute* in English cannot be followed by an infinitive [line 16].

The last correction was a VSO construction that could not easily be 'quick-fixed' (see Vasconcellos 1986a) and required the movement of four words to the end of the sentence.

Of the nine syntactic corrections, two (22%) were made using macros, indicating that these were operations commonly performed by posteditors. (Use of a macro, of course, speeds up the process.)

Cohesive devices. Of the 21 cohesive devices, 12 (57%) had to do with definiteness, a subcategory of referential cohesion: two called for insertion of the definite article and nine for its deletion (the changes on line 2 were counted twice, once as deletion of the article and once as conjunction). Seven of the other nine devices could be accounted for in terms of conjunction. In five instances, conjoining of the terms in an enumeration was highlighted cohesively by repetition of the preposition [lines 2 (twice), 5, 6] or downgraded by the deletion thereof [line 20]. In another case [line 9], a relative clause marker was changed from a comma to a dash, giving more independence to the conjoining relation. Also under conjunction, the head noun *capacity* was redundant in the premodifying enumeration of the NP whose head was *performance*. The other two changes had to do with discourse texture. In one, the information structure was preserved by postponing the concept *disadvantaged* [line 7] after *populations*. In the other, movement of the word *usually* (from the Spanish *generalmente*) to the front of the clause gave it thematic status [line 10].

Of the 21 cohesive devices, 12 (57%) were introduced using macros.

There were also interpretations for coherence at this level. The word *exist* [line 7] emphasizes the notion of 'existence' in a context where it does not apply. Changing the translation *usually* to *in general* [line 10, counted previously as a move for purposes of thematization] brackets the clause that follows and appears to approximate more closely the meaning originally intended. Finally, deletion of *aspects of* seems to tighten the coherence in English.

4.2 Final translation. At this maximum level of refinement 14 additional changes were made, which showed the following breakdown:

Syntactic corrections	0	0%
Cohesive devices	(?) 3	21%
Interpretations for coherence	11	<u>79%</u>
		100%

As it can be seen, there was a clear preponderance of interpretations with a view to improving coherence. Only three of the changes could be regarded as cohesive devices, and in each case an underlying motivation of coherence could be argued.

One of the changes that was classified as cohesive was the replacement of *however* by *nevertheless* [line 3]. In surface structure, this is a cohesive relation expressed through a conjunction. On the other hand, the interpretation that led to the change might well be considered to involve coherence. This was also true of the changes in the conjunctions from *in*

general to *in sum* [line 10]. The other candidate cohesive device was the use of commas to bracket the phrase on lines 24-25. Again, although the device is a surface-structure mechanism, one could argue that it was necessitated by the expansion which had been added between the commas.

Eight (73%) of the remaining 11 changes were clear-cut expansions beyond the propositional content given by the machine translation [two insertions on line 7 plus those on lines 16, 17, 18, 24, 26, and 27]. Because the material had not been present in the discourse, these changes could only be classed as interpretations for coherence. Two others [lines 10 and 40], although they did not expand the number of words in the text, added further semantic specificity which had not been there before, and in this sense they were also expansions. Finally, the use of *to* [line 28] adds force to the claim being made.

5 Discussion. The foregoing analysis bears out the difficulty of separating cohesion and coherence. In several of the examples it seemed that even though cohesive devices had been used, because of the circumstances of postediting there was also a strong component of interpretation for coherence.

For instance, the changing of *however* to *nevertheless* [line 3] and *in general* to *in sum* [line 10] were both further refinements of cohesive relations that were already present in the discourse--and in fact had been introduced during the first pass. It looks as if two different types of motivation were at work. In the first pass, the need for a cohesive tie was detected, and the material introduced was a close approximation of the original Spanish. In the final polishing, however, the posteditor became interpreter and proceeded to introduce semantic components which represented a slight departure from the unmarked meaning of these conjunctions, doing so in the interest of coherence.

What have we learned from this exercise? In the sample studied it was clear that syntactic corrections and cohesive devices predominated in the first pass and that interpretations for coherence accounted for the changes in the final translation.

In the haste of work, the distinction between these two levels tends to blur: during the first pass it may happen that interpretations are introduced, while in the final review action may be taken on opportunities that were missed the first time around. It is not reasonable to expect that posteditors will follow a rigorous separation between the two. Still, time can be saved for some applications if an effort is made to limit changes to syntactic corrections and cohesive devices.

As far as the contribution to MT development is concerned, it is reasonable to hope that many cohesive devices can eventually be written into basic algorithms or inter- or postprocessors. On the other hand, it is also important to recognize the posteditor's role as interpreter of coherence, and to understand that this aspect of human performance is beyond formalization.

Una nutrición adecuada es esencial para la salud del individuo, la productividad colectiva y el bienestar social. Sin embargo, las carencias nutricionales continúan altamente prevalentes en la Región, particularmente la desnutrición energético-proteica y las deficiencias de hierro, de vitamina A y de yodo.

Es evidente que existen poblaciones subprivilegiadas en términos de disponibilidad y consumo de alimentos, atención de salud, saneamiento ambiental, educación, oportunidades de trabajo y organización social, las cuales viven generalmente en situación de pobreza crítica. Dichas condiciones están agravadas por los cambios demográficos que están ocurriendo en Latinoamérica y el Caribe y que crean nuevas demandas al sistema alimentario.

Las deficiencias nutricionales agravan los problemas de salud y contribuyen a elevar las tasas de morbilidad y mortalidad especialmente en niños menores de 5 años, ocasionando alteraciones funcionales con efectos inmediatos y repercusiones a largo plazo en las áreas de capacidad mental, social, inmunológica, reproductiva y de rendimiento físico.

OK / Adequate nutrition is essential for health of the individual, the collective productivity and the social well-being. However, the nutritional deficiencies continue to be highly prevalent in the Region, particularly the +energy-protein malnutrition+ and deficiencies of iron, of vitamin A and of iodine.

PP 21 TU 7 It is evident that there exist disadvantaged populations in terms of availability and consumption of food, health care, environmental sanitation, education, job opportunities and social organization, which live usually PP 17 TU // in a state of critical poverty. Said conditions are aggravated by the demographic changes that are occurring in Latin America and the Caribbean and that place new demands to the food system.

PP 45 TU 15 Nutritional deficiencies aggravate the health problems and contribute to increase the rates of morbidity and mortality especially in children under 5, causing functional alterations with immediate effects and long-term repercussions in the areas of mental capacity, social, +immunological+, reproductive and of physical performance.

Concomitantemente, otros grupos de población sufren enfermedades crónicas—cardiovasculares, diabetes no insulino dependiente, obesidad y algunos tipos de cáncer—en cuya etiología juega un papel importante la malnutrición debida a imbalance de nutrientes y exceso relativo de energía.

Puesto que la nutrición y la salud no pueden existir en forma independiente y ambas son esenciales para el desarrollo y el bienestar de la población, una alimentación adecuada es esencial para ambos aspectos del bienestar individual y colectivo. Todos los factores que inciden en la disponibilidad, adquisición, distribución intrafamiliar, consumo y utilización de los alimentos deben tenerse en cuenta en los planes y programas dirigidos a lograr y mantener un buen estado de salud y nutrición.

El control y prevención de la malnutrición no puede ser responsabilidad exclusiva del sector salud. Es necesario poner en marcha políticas y programas coordinados dirigidos a la identificación, vigilancia y corrección de los diversos factores que afectan el estado nutricional y el consumo de alimentos.

NO 23 2/ At the same time, other population groups suffer chronic 22 diseases—cardiovascular, non-insulin dependent diabetes, 23 obesity and some types of cancer—in whose etiology plays 24 an important role malnutrition due to imbalance of 25 nutrients and relative excess of energy.

NO 14 TU 26 Because nutrition and health cannot exist independently 27 and both are essential for the development and the well- 28 being of the population, an adequate diet is essential for 29 both aspects of the individual and collective well-being. 30 All the factors that affect the availability, acquisition, 31 intrafamily distribution, consumption and utilization of 32 the food should be taken into account in the plans and 33 programs directed toward achieving and maintaining a good 34 state of health and nutrition.

OK 35 The control and prevention of malnutrition cannot be 36 exclusive responsibility of the health sector. It is 37 necessary to implement policies and coordinated programs 38 directed toward the identification, surveillance and 39 correction of the various factors that affect the 40 nutritional status and the consumption of food.

Appendix B: Target-only raw MT output with first-pass changes.

- 1 Adequate nutrition is essential for ~~the~~ health of the ~~for~~
- 2 individual, ~~the~~ collective productivity and ~~the~~ social well-being. However, ~~the~~ nutritional deficiencies continue to
- 4 be highly prevalent in the Region, particularly ~~the~~ + energy-
- 5 protein + malnutrition and deficiencies of iron, ~~of~~ vitamin A
- 6 and ~~of~~ iodine.
- 7 It is evident that there ~~exist~~ are disadvantaged populations, ~~which are~~
- 8 in terms of availability and consumption of food, health
- 9 care, environmental sanitation, education, job
- 10 opportunities and social organization, ~~which~~ live ~~usually~~ ^{-in general,}
- 11 in a state of critical poverty. Said conditions are
- 12 aggravated by the demographic changes that are occurring in
- 13 Latin America and the Caribbean, ~~and~~ ^{which} place new demands
- 14 ~~on~~ ^{on} the food system.
- 15 Nutritional deficiencies aggravate ~~the~~ health problems
- 16 and contribute to increase ~~the~~ ^d rates of morbidity and
- 17 mortality, especially in children under 5, causing
- 18 functional alterations with immediate effects and long-term
- 19 repercussions in the areas of mental ~~capacity~~, social,
- 20 + immunological +, reproductive and ~~of~~ physical performance.
- 21 At the same time, other population groups suffer ~~from~~ chronic
- 22 diseases--cardiovascular, non-insulin-dependent diabetes,
- 23 obesity and some types of cancer--in whose etiology ~~plays~~
- 24 an important role malnutrition due to imbalance of

- 25 nutrients and relative excess of energy.
- 26 Because nutrition and health cannot exist independently
- 27 and both are essential for the development and the well-
- 28 being of the population, an adequate diet is essential for
- 29 both aspects of the individual and collective well-being.
- 30 All the factors that affect the availability, acquisition
- 31 intrafamily distribution, consumption and utilization of
- 32 the food should be taken into account in the plans and
- 33 programs directed toward achieving and maintaining a good
- 34 state of health and nutrition.
- 35 The control and prevention of malnutrition cannot be the
- 36 exclusive responsibility of the health sector. It is
- 37 necessary to implement policies and coordinated programs
- 38 directed toward the identification, surveillance and
- 39 correction of the various factors that affect the
- 40 nutritional status and the consumption of food.

Appendix C: First-pass postedit with final changes.

- 1 Adequate nutrition is essential for the health of the
- 2 individual, for collective productivity and for social well-
- 3 being. *Nevertheless*, *However*, nutritional deficiencies continue to
- 4 be highly prevalent in the Region, particularly energy-
- 5 protein malnutrition and deficiencies of iron, vitamin A
- 6 and iodine.
- 7 *Unquestionably*, *It is evident* that there are *some* populations which are severely disadvantaged

- 8 in terms of availability and consumption of food, health
- 9 care, environmental sanitation, education, job
- 10 opportunities and social organization--in ^{sum} general, which live
- 11 in a state of critical poverty. These conditions are
- 12 aggravated by the demographic changes that are occurring in
- 13 Latin America and the Caribbean, which place new demands
- 14 on the food system.
- 15 Nutritional deficiencies aggravate health problems
thus
- 16 and contribute to increased rates of morbidity and
- 17 mortality, especially in children under 5, ^{as well as} causing
that have both
- 18 functional alterations with immediate effects and long-term
- 19 repercussions in the areas of mental, social,
- 20 immunological, reproductive and physical performance.
- 21 At the same time, other population groups suffer from chronic
- 22 diseases--cardiovascular, non-insulin-dependent diabetes,
- 23 obesity and some types of cancer--in whose etiology
- 24 malnutrition ^{in the form of} due to ^{imbalance of}
intake
- 25 nutrients and relative excess energy, plays an important role.
- 26 Because nutrition and health cannot exist ^{one} independently
of the other together they are
- 27 and both are essential for the development and the well-
- 28 being of the population, an adequate diet is essential ^{to}
- 29 both individual and collective well-being.
- 30 All the factors that affect the availability, acquisition
- 31 intrafamily distribution, consumption and utilization of

- 32 food should be taken into account in the plans and
- 33 programs directed toward achieving and maintaining a good
- 34 state of health and nutrition.
- 35 The control and prevention of malnutrition cannot be the
- 36 exclusive responsibility of the health sector. It is
- 37 necessary to implement policies and coordinated programs
- 38 directed toward identification, surveillance and
- 39 correction of the various factors that affect
- 40 nutritional status and ~~the consumption of food~~^{intake}.

Appendix D: Final translation.

- 1 Adequate nutrition is essential for the health of the
- 2 individual, for collective productivity and for social well-
- 3 being. Nevertheless, nutritional deficiencies continue to
- 4 be highly prevalent in the Region, particularly energy-
- 5 protein malnutrition and deficiencies of iron, vitamin A
- 6 and iodine.
- 7 Unquestionably, there are some populations which are severely
- 8 disadvantaged
- 9 in terms of availability and consumption of food, health
- 10 care, environmental sanitation, education, job
- 11 opportunities and social organization--in sum, which live
- 12 in a state of critical poverty. These conditions are
- 13 aggravated by the demographic changes that are occurring
- 14 in Latin America and the Caribbean, which place new demands
- 15 on the food system.
- 16 Nutritional deficiencies aggravate health problems
- 17 and thus contribute to increased rates of morbidity and
- 18 mortality, especially in children under 5, as well as causing
- 19 functional alterations that have both immediate effects and long-term
- 20 repercussions in the areas of mental, social,
- 21 immunological, reproductive and physical performance.
- 22 At the same time, other population groups suffer from chronic
- 23 diseases--cardiovascular, non-insulin dependent diabetes,
- 24 obesity and some types of cancer--in whose etiology
- 25 malnutrition, in the form of unbalanced
- 26 nutrient intake and relative excess energy, plays an important role.
- Because nutrition and health cannot exist one independently of the

27 other and together they are both essential for the development and the well-being of the population--an adequate diet is essential
 29 both to individual and to collective well-being.
 30 All the factors that affect the availability, acquisition
 31 intrafamily distribution, consumption and utilization of
 32 food should be taken into account in the plans and
 33 programs aimed at achieving and maintaining a good
 34 state of health and nutrition.
 35 The control and prevention of malnutrition cannot be the
 36 exclusive responsibility of the health sector. It is
 37 necessary to implement policies and coordinated programs
 38 aimed at identification, surveillance and
 39 correction of the various factors that affect
 40 nutritional status and food intake.

Notes

1. This is in fact the rationale behind the Augmentor component of KBMT-89, the knowledge-based MT constellation being developed at Carnegie Mellon University (Nirenburg in this volume).

2. Löfller-Laurian's general guidelines for conventional postediting (1986, translation from the French by MV): respect the raw translation as much as possible; change only that which absolutely must be changed; make the changes as simple as possible; and change what is unfaithful, incorrect, or incomprehensible. Her specific rules are: (1) provide the correct technical terms, proper names, and abbreviations; (2) resolve ambiguities; (3) check relationships between verbs and their arguments and within NPs; (4) check logical relationships in long sentences; (5) when restructuring is necessary, chose the most economical approach; (6) watch for differences in punctuation between the two languages; (7) watch for differences in verb tenses between the two languages; (8) change modality and qualification to conform to target language usage; (9) make certain that negations are correctly rendered; (10) impose parallel structure in enumerations; (11) provide functional equivalents for idiomatic phrases; and finally, (12) concentrate on going straight to the point. She concludes by emphasizing that to work quickly does not mean that quality has to be sacrificed.

3. It will be seen later below that in the case of cohesion this interpretation fits with the model of Halliday and Hasan (1976) and in the case of coherence with that of Sanders (1987).

4. Semantic meanings are given between single quote-marks, whereas English lexical items are italicized.

5. Types of meaning other than propositional meaning are variously categorized in different linguistic models. Bühler (1934) identified three functions of language--the referential, the expressive, and the conative. Jakobson (1960) developed a scheme of six. Halliday (1977) sees language as having three generalized functions--ideational, interpersonal, and textual--each of which corresponds to a subset of interdependent systems that convey different types of meaning. Sanders (1987) speaks of propositional

content, illocutionary act, and implicature. Common to all these models is the fact that, depending on the intentions of the communicator, the choice can be made to focus on one or another of the meaning systems while the remaining systems still continue to be instantiated in the text.

6. See Brown and Yule (1983:195-199) for a discussion of the place of semantic connectedness.

7. See Stubbs (1983:84-103) for a discussion of discoursal well-formedness.

8. SPANAM™ (Spanish-English) and ENGSPAN™ (English-Spanish) are MT systems developed in-house by the Pan American Health Organization in order to meet internal translation needs (see Vasconcellos and León 1988). SPANAM has been in practical use since January 1980 and ENGSPAN (development partially supported by Grant DPE-5543-G-SS-3048-00 from the U.S. Agency for International Development), since 1985. ENGSPAN has since been installed at AID and at international agricultural research centers in Colombia and the Philippines. SPANAM's dictionaries have more than 62,000 entries; ENGSPAN's, about 55,000.

9. Segmentation of the sample was dictated by the amount of text that would fit on a two-page side-by-side display.

10. These calculations do not include the serial comma, which is house style, or deletion of the plus (+) signs indicating terminological reliability. For both these operations there are macros which accomplish them quickly.

11. The sample text had been purged of dictionary problems that could easily be remedied, and in this sense the experience was different from typical postediting, which would be expected to include more corrections for such basic errors as not-found words.

References

- Bostad, Dale A. 1987. Machine translation: The USAF experience. *Proceedings of the 28th Annual Conference of the American Translators Association* (Cleveland, 16-19 October 1986), ed. by Karl Kummer. 435-43. Medford, N.J.: Learned Information, Inc.
- Brown, Gillian, and George Yule. 1983. Discourse analysis. Cambridge, London: Cambridge University Press.
- Bühler, Karl. 1933 [1982]. The axiomatization of the language sciences. *Translation of Die Axiomatik der Sprachwissenschaften* by Robert E. Innis. In: Karl Bühler: Semiotic foundations of language theory. 77-164. New York and London: Plenum.
- Halliday, M.A.K. 1967. Notes on transitivity and theme in English. Parts 1 and 2, *Journal of Linguistics* 3.37-81, 199-244. Part 3, 4.179-215.
- Halliday, M.A.K. 1977. Text as semantic choice in social contexts. Grammars and descriptions, ed. Teun A. van Dijk and Janos Petöfi. Berlin and New York: De Gruyter. 176-225.
- Halliday, M.A.K., and Ruqaiya Hasan. 1976. Cohesion in English. London: Longman.
- Jakobson, Roman. 1960. Closing statement: Linguistics and poetics. *Style in language*, ed. by Thomas A. Sebeok. 350-77. Cambridge, Mass.: MIT Press; New York and London: Wiley.
- Löffler-Laurian, Anne-Marie. 1986. Post-édition rapide et post-édition conventionnelle. Part 1, *Multilingua* 5.81-8. Part 2, 5.225-29.
- Newman, Patricia E. 1988. Information-only machine translation: A feasibility study. In: *Technology as translation strategy*, ed. by Muriel Vasconcellos. 178-189. Binghamton: State University of New York Press. American Translators Association Scholarly Monograph 2.
- Sanders, Robert E. 1987. Cognitive foundations of calculated speech. Albany: State University of New York Press.

- Schiffrin, Deborah. 1987. Discourse markers. Cambridge, New York: Cambridge University Press. *Studies in Interactional Sociolinguistics* 5.
- Stubbs, Michael. 1983. Discourse analysis: The sociolinguistic analysis of natural language. Chicago: University of Chicago Press.
- Van Dijk, T.A. 1972. Some aspects of text grammars. The Hague: Mouton.
- Vasconcellos, Muriel. 1985. Theme and focus: Cross-language comparison via translations from extended discourse. Georgetown University dissertation.
- Vasconcellos, Muriel. 1986a. Functional considerations in the postediting of machine-translated output: Dealing with V(S)O versus SVO. *Computers and Translation* 1.21-38.
- Vasconcellos, Muriel. 1986b. Humor through the listener's voice: A functional model for the capture of humor in translation. *Babel* 32.134-45.
- Vasconcellos, Muriel, and Marjorie León. 1988. SPANAM and ENGSPAN: Machine translation at the Pan American Health Organization. In: *Machine translation systems*, ed. Jonathan Slocum. 187-235. Cambridge, New York: Cambridge University Press.
- Wagner, Elizabeth. 1985. Rapid post-editing of Systran. In: *Tools for the trade: Translating and the Computer* 5, ed. by Veronica Lawson. 199-213. Amsterdam: North-Holland.
- Widdowson, H.G. 1979. Explorations in applied linguistics. London and Edinburgh: Oxford University Press.

Second language acquisition: Do we really want a unified theory?

Richard Lutz
Georgetown University

1 Introduction. The subjects covered by the heading *Language Teaching, Testing, and Technology* do not explicitly include one of the greatest shifts in emphasis to take place in the area of applied linguistics since *GURT* began: that shift is toward the language learner and the processes involved in language acquisition. But the omission tells us much about how far second language acquisition (SLA) research has progressed in recent years. We now assume, for example, that the teaching process is actually a kind of 'facilitation of learning' (Brown 1987:1). We now know enough about SLA for those in related applied fields to rely on our work for the basic theoretical underpinnings of language teaching, testing and technology. It is the nature of those theoretical underpinnings that I will discuss below.

The title of this paper may be a bit misleading. It suggests that the author has looked at the field of Second Language Acquisition, assessed its needs and capabilities, and reached firm conclusions. On the contrary, this yes/no question is just that: a question, and one which really requires a kind of consensus approach that neither I nor the dominant figures in the field of SLA research can reach independently of one another. Indeed, it is the group work of all those involved in the field of SLA which decides the answers to such questions. This paper will consider only a few aspects of the question, and by its conclusion, the reader may decide that there are more fitting questions raised in this paper than that of our wishes. Those questions must include whether a unified theory is possible or likely, as well as what benefits and deficits such an approach to SLA may bring to the field of applied linguistics.

There appears to be something of a consensus at present, both among researchers in SLA itself, and those involved in applied fields involving language pedagogy, that a unified theory of language acquisition is desirable or even that it is essential. Brown (1987:240-41) assumes the goal of SLA research to be, ultimately, a 'full theory' of how language is acquired. The very complexity of SLA, he states,

means that there are so many separate but interrelated factors within one intricate entity that it is exceedingly difficult to bring order and simplicity into that entity...We must nevertheless pursue the task of theory building.

Hammerly (1985:xi) blames the lack of a unified theory (in the context of language teaching) for a variety of applied maladies:

The lack of an adequate theory affects all our activities. Whole series of books on language teaching are being published, by several publishers, without theories to which the content of the books can be related. What research there is can only be haphazard without an integrated theory to guide it--almost all theories, past and present, have been partial in nature, emphasizing some facts and ignoring the rest.

Hammerly goes on to blame this 'theoretical vacuum' for 'the confusion and faddism' in the field of language teaching:

A commitment to excellence means, among other things, refusing to 'go with the flow' unless it happens to be flowing in the right direction. But how can people know what the 'right direction' is without an adequate, explicit, comprehensive theory?

It would be misleading, however, to present SLA researchers as united behind the a priori, deductive 'theory-then-research' (Reynolds 1971) approach. Schouten (1979, cited in Ellis 1985) argues that a proliferation of models in SLA has actually stifled research. An inductive 'research-then-theory' approach, unconstrained by the preconceptions of an extant theory, may delve into important aspects of the process that would otherwise go unnoticed. At any given point in time, 'theory-first' deductivists (following Popper 1959) will be able to identify which data are relevant and which experiments need to be undertaken, and, I might add, the theory-first approach could end up saving a lot of wasted time and energy. Kuhn (1962:4-5) suggests that the concept of 'development-by-accumulation' itself is probably mistaken, since, as more and more research accumulates, it becomes increasingly clear that the individual researchers were working within a paradigm all the time, that is, that the scientific community must have a shared set of assumptions about what the world is really like. Whether these assumptions are made explicit, researchers are none the less tradition-bound.

The epigraph of a recent book on language teaching quotes Bernard le B. de Fontenelle as follows:

To despise theory is to have the excessively vain pretension to do without knowing what one does, and to speak without knowing what one says.
(Hammerly 1985)

But if there is disagreement about which path (theory-first or research-first) will lead to an integrated view of the SLA world most efficiently (and most would probably assume that both approaches are viable and indeed needed), there seems little disagreement about the ultimate goal: understanding SLA as a monolithic phenomenon.

Let it be clear from the onset that the value of theory-building, and of specific theories--that is, testable hypotheses--in particular, is not in doubt. Theories are really just 'complexes of hypotheses' (Hammerly 1985:xii). Researchers must investigate hypotheses rigorously, and certain corroborated hypotheses will naturally lead to others. The present concern is one of an

integrated view, one which would serve to guide all those interested in the larger questions of HOW and WHY we acquire language.

Hakuta (1981:1) identifies the game of language acquisition research (both first and second) as 'the search for an appropriate level of description of the learner's system of rules.' Descriptive adequacy is certainly one of the minimal requirements of any theory, but is unlikely to satisfy our desire to understand how and why language acquisition takes place. A unified theory would have to answer these questions, and do so in a manner that reassures us that we are asking the right questions, and accounting for all the factors that really matter. Hakuta's use of the term 'appropriate level' strikes me as critical, and this topic is the focus of section 2.

2 Scope of the inquiry. Before one can approach the question of whether a unified acquisition theory is possible, the boundaries which define the field of inquiry must be marked. If the question to be addressed is how second languages are acquired, we must understand what we mean by language, by learning, and, in many contexts, by teaching. The purpose of a theoretical paradigm has been expounded upon most thoroughly by Kuhn (1962). It is the paradigm which sets the limits of legitimate inquiry. Widdowson (1988:185) discusses the relevance of PARADIGM to the field of applied linguistics as follows:

[C]ertain topics are immediately recognized as self-evidently relevant: integrative tests, comprehensible input, relative clauses, group work. Other topics are, just as readily, dismissed as irrelevant: astronomy, gastronomy, the imagery of *Macbeth*, the erotic verse of the Earl of Rochester.

The paradigm, then, serves to demarcate: this is a necessary and beneficial aspect of working as part of the scientific community toward a certain goal. There is a hazard to this approach, however. Tannen (1989:11), in a discussion of the place of discourse analysis among the subfields of linguistics, finds the interdisciplinary character of discourse analysis to be a source of criticism by many. Like all interdisciplinary endeavors, it can be viewed as too heterogeneous, lacking in a monolithic theory, which may lead to charges of intellectual 'immaturity'. She notes Widdowson (185-86), who writes:

This means that those who try to promote cross-cultural relations by being inter-disciplinary are likely to be ostracized by both sides and to be stigmatized twice over as amateur and mountebank.

Tannen (1989:13) also warns that this interdisciplinary quality does not 'exempt individual works (or individuals' work) from having to make clear theoretical, methodological, and, when appropriate, empirical frameworks.' Thus, as I indicated above, the issue is not one of whether researchers are to work from theoretical perspectives, but whether all research should be tied together to form one metatheory.

A description of SLA reveals the extent of its interdisciplinary nature. It is firmly a part of linguistics, psychology, neurology, biology, sociology and

education (and I will expect readers to throw in other disciplines which I have neglected). To present a unified front, SLA researchers would have to be nimble indeed. A paradigm sets limits: but even those pursuing an integrated theory of SLA are obliged to check their findings with those in related areas. Criticisms of a variety of theories and models of SLA have invariably drawn on other banks of knowledge. Spolsky (1985:281), in his search for a unified theory of SLA, writes repeatedly of the negative effects of this isolation:

As linguists often tend to forget, learning theory is the special area of psychology. Lulled by our belief that Chomsky overthrew Skinner, who had earlier cast aside Pavlov, we have been trying to build our own models of learning, and the results of amateur work appear. But it is surely to be expected that there would be psychologists who have tried not to abandon but, in the traditional way of all paradigms, patch old models by seeing what they can incorporate of the new.

The great danger of a discrete or closed universe of SLA is that it will be maintained in isolation from research and ideas which impinge quite clearly on its own world view.

At present, most models of SLA display an intellectual bias in favor of one or two related disciplines. Lamendella (1979) draws on neurolinguistic research, looking to the anatomy of the brain for information about language processing. Schumann's (1978) Acculturation Model eyes sociological and other cultural variables as primary, and individual affective variables as secondary but related. Andersen (1980) modifies Schumann's model to emphasize the role that individual cognitive variables play in SLA. Hatch (1983) has looked at SLA from the perspective of discourse and how social interaction via discourse controls SLA development. Certainly the most prominent model of SLA is presented in the work of Krashen (1980, 1982, etc.), referred to by Spolsky (1985: 270) as the 'Extended Monitor theory', and by Krashen (personal communication) most recently as the 'Input theory'. It is the most comprehensive of the available models, but that comprehensive quality is achieved largely by combining a variety of smaller, relatively discrete models. His Affective Filter corresponds to parts of Schumann's Acculturation Model, to explain how far (and how fast) the L2 user proceeds toward the target goal (i.e., intake becoming input). His Monitor is used to account for discrepancies found in the performance of the L2 user (i.e. output). While the Input theory can be commended for covering a variety of aspects of SLA, there is no obvious (unifying) principle connecting all the parts of the overall theory.¹ Such comprehensive theories may be possible to describe, but there are limits to how integrated they can be.

3 Integrating parts. Physicists have been searching for decades for the fundamental forces in nature. Since Einstein, the search has been on for the Grand Unified Theory that will unite gravity, electromagnetism, the strong nuclear force and the weak nuclear force. The theory will explain 'everything', at least 'everything' that comes within the purview of physics. Stephen Hawking, a leader of this quest, has been quoted as stating his goal to be 'a complete understanding of the universe, why it is as it is and why it exists at

all' (Allen: 1988). The search for a unified theory (known in some circles as 'TOE'--the theory of everything) was precipitated by conflicts between two leading theories in the 1940s, Einstein's theory of relativity, which dealt with large bodies and gravity--the universe as an expanding balloon--and quantum mechanics, which explained the behavior of forces at the level of the atom and smaller. Quantum mechanics, in turn, has discovered that all the order of the universe rests on a foundation of chaos (i.e. violent, random motion) at the particle level. Without pretending to understand the details, I gather that the conflict is between the apparent predictability at the level of the expanding universe and the chaos at its base. To resolve the conflict, Hawking has relied on those features of quantum mechanics which are least in doubt, and applied them in a kind of 'what if' scenario to the larger cosmos. The result is a unified model of a universe forever developing bulges on its sides, which give birth to new 'baby' universes, connected by momentary passageways or 'wormholes'. Hawking contends that his model of an infinity of baby universes 'is fairly independent of the details of the final theory' (Waldrop 1989). The final explanation will have to resemble his theory since relativity and quantum mechanics each conform to the observable world. The two have to be integrated.

What would a comparable 'TOE' for SLA look like? It would not be a comprehensive collection of separate and relatively unconnected parts. Rather, it would need to integrate all the known forces at work in SLA: it would need to predict the result of any given combination of conditions. Spolsky has proposed a 'preference' model (Jackendoff 1983:157) which can 'derive a quasi-determinate result from unreliable data.' The outcome of learning a second language would depend on sets of (a) necessary conditions, (b) graded conditions in which there is a relation between the extent to which a condition is met and the outcome, and (c) typicality conditions that typically apply, but may not (Spolsky 1985: 282). Whether such a complex model is feasible or even testable is far from evident. As I will argue below, even if it were possible to construct a grand unified theory using something akin to the preference system, it is questionable whether a theory which covers everything is the most desirable theory. Unlike Einstein's 'simple universe', described by a unifying force--energy--(Calder 1979:207ff.), the preference system assumes SLA to be the complex interaction of a huge number of factors of varying importance.

4 Unification and dichotomies. Are all the various variables that would make up the grand unified theory of SLA really of a kind? That is, do they belong in the same paradigm?

Acquisition theory is replete with dichotomies: nature and nurture, learning and acquisition, learning and teaching. It is clear that some of these pairs belong together, as complements to one another or as end points on a continuum. In recent years, SLA researchers have gone a long way toward reconciling the differences between the nativist and behavioral traditions by arguing that it is the INTERACTION between biology and the environment which best accounts for the observed phenomena. Some have attacked Krashen's strong noninterface position that acquisition and learning are entirely separate: McLaughlin (1978) redefined them in terms of the degree to which the

processing of information is controlled or automatic, and in terms of whether a given behavior receives focal or peripheral attention. Clearly, dichotomies as such are not inconsistent with a unified approach, just as Einstein's energy and matter proved to be equivalent. They belong in the same paradigm. But not all such pairs can be equated. Evidence from first language acquisition research suggests that comprehension and production should NOT be treated as part of one process: there are, for instance, considerable differences in the way children extend words in comprehension and in production (Fremgen and Fay 1980, Kuczaj 1979, Leonard 1983, cited in Pease, Berko Gleason and Pan 1989). The two systems do not really overlap: rather, each develops rather independently of the other, and to the many known incidences of a child understanding something but unable to express it, there are reported cases of children using a word accurately when comprehension checks make it clear that they don't understand that same word when others use it (Rice 1983). While no one would claim that comprehension and production are unrelated phenomena, these data do suggest that the processes involved in each are sufficiently different to require separate treatment.²

Conversely, a unified view of SLA must necessarily draw the line between SLA and other, related phenomena. Most notable among them is first language acquisition. Entire theories have been promulgated on the question of whether SLA and FLA are part of the same process, or should be treated differently. It should be the case that any theory which assumes the two to be closely related will need a unified theory not of SLA but of acquisition generally. Why not then unified theories of ALL acquisition?

Similarly, numerous researchers approach acquisition as part of general human learning and cognition. A true unified theory for them would set the boundaries far into the field of cognitive psychology, and the distinction between verbal learning (both first and second) and nonverbal learning would be deemphasized. Larsen-Freeman (1983) surveyed four leading models of SLA and pronounced each of them lacking the essential cognitive component. But on the other hand, as Larsen-Freeman would say, language acquisition is more than simply 'verbal learning', and certainly involves key factors which do not fit neatly into the realm of 'cognition.'

The tension then is between a need to look at ever expanding areas of human experience to achieve a holistic view of the phenomenon of acquisition ('baby universes' all connected together through 'wormholes' of interdisciplinary data), and the need to recognize the independence of various levels of the acquisition universe.

5 Practical issues and conclusions. Many involved in language pedagogy will be sympathetic to Hammerly's (1985) lament that without a unified theory of language TEACHING, the classroom is subject to the whims of faddists. He is very clear, however, to separate the classroom from 'natural' settings: however valid the theories of SLA may be for the 'real world', a special set of conditions obtain in the classroom, and a special, that is, separate, theory of teaching is required. He clearly understands the 'multidisciplinary' (111) nature at work when he writes of the contribution that psychology has to play in most language learning activities. But, he writes, 'the choice of "applicable" psychological theories and concepts is up to us [viz. pedagogues]' (81)

Hammerly is really arguing that SLA researchers, sociolinguists and psychologists, and others should leave language teaching alone, and let those involved in that specialized field of 'linguistics' (ix and throughout) see what contributions from those related fields make sense in a classroom context, then go about testing methodologies which incorporate those ideas. It should be eclectic, drawing on the best that others have to offer. In a sense, those of us in SLA research should heave a collective sigh of relief: the burden has been removed from us. For historical and practical considerations, SLA has been considered as a stepping-stone to teaching methodology and other applied fields, and I suspect that this obligation has provided much of the impetus for a unified theory. Looked to for the answer to the question of how to teach, SLA researchers have had to view language learning as a unitary phenomenon. The resolution to the issue of how integrated the field of SLA needs to be depends on what we want to accomplish. We are stuck with a huge if not infinite number of 'baby universes' in SLA, all contributing to our understanding of how human beings acquire language. Any unified theory is bound to neglect at least some of those universes. What we have at present--essentially models that look at the contribution of one or more factors--may well be the best (or at least the most desirable) of all possible worlds: one which promotes contributions from a variety of frameworks, with ever-expanding horizons.

Notes

1. Krashen (personal communication) objects to my analysis, claiming input to be the unifying feature of his theory. It appears to me that Krashen's insistence on the dichotomous nature of learning versus acquisition belies any claims to unity: he is essentially claiming that two processes do NOT belong to the same paradigm. Learning and the Monitor are useful to him precisely because they allow him to exclude that which doesn't belong in the acquisition paradigm.

2. Rice (1983:353) cites numerous studies which claim to support a 'weak cognition' view (i.e. cognition accounts for some, but not all aspects of acquisition). She cites examples of language acquisition as independent of meanings, including children who learn new, more complicated ways of expressing the same meaning (e.g. shifts from proper names to deictic terms).

References

- Allen, H. 1988. The mind in pursuit of the universe: Physicist Stephen Hawking and the never-answered question. *Washington Post*, April 15, B1.
- Andersen, R. 1980. The role of creolization in Schumann's pidginization hypothesis for second language acquisition. In: *Research in second language acquisition*, ed. R. Scarcella and S. Krashen. Rowley, Mass.: Newbury House.
- Berko Gleason, J., ed. 1989. *The development of language*. 2nd ed. Columbus, Ohio: Merrill.
- Brown, H.D. 1987. *Principles of language learning and teaching*. 2nd ed. Englewood Cliffs, N.J.: Prentice-Hall.
- Calder, N. 1979. *Einstein's universe*. Harmondsworth, U.K.: Penguin.
- Ellis, R. 1985. *Understanding second language acquisition*. Oxford: Oxford University Press.

- Fremgen, A., and D. Fay. 1980. Overextension in production and comprehension: A methodological clarification. *Journal of Child Language* 7:205-11.
- Gingras, R., ed. 1978. Second language acquisition and foreign language teaching. Arlington, Va.: Center for Applied Linguistics.
- Hakuta, K. 1981. Some common goals for second and first language acquisition research. In: *New dimensions in second language acquisition research*, ed. R. Andersen. Rowley, Mass.: Newbury House.
- Hammerly, H. 1985. An integrated theory of language teaching and its practical consequences. Blaine, Wash.: Second Language Publications. Vol. 2 of the Series in Linguistics.
- Hatch, E. 1983. Psycholinguistics: A second language perspective. Rowley, Mass.: Newbury House.
- Jackendoff, R. 1983. Semantics and cognition. Cambridge, Mass.: MIT Press.
- Krashen, S. 1980. The theoretical and practical relevance of simple codes in second language acquisition. In: *Research in second language acquisition*, ed. R. Scarcella and S. Krashen. Rowley, Mass.: Newbury House.
- Krashen, S. 1982. Principles and practice in second language acquisition. Oxford: Pergamon Institute of English.
- Kuczaj, S.A. 1979. Young children's overextensions of object words in comprehension and/or production: Support for a prototype theory of early word meaning. Paper presented at the meeting of the Society for Research in Child Development. San Francisco.
- Kuhn, T.S. 1962. The structure of scientific revolutions. Chicago, Ill.: University of Chicago Press.
- Lamendella, J. 1977. General principles of neurofunctional organization and their manifestation in primary and nonprimary language acquisition. *Language Learning* 27: 155-98.
- Larsen-Freeman, D. 1983. Second language acquisition: Getting the whole picture. In: *Second language acquisition studies*, ed. K.M. Bailey, M.H. Long, and S. Peck. Rowley, Mass.: Newbury House.
- Leonard, L.B. 1983. Production before comprehension: Some methodological issues. Paper presented at the Society for Research in Child Development, Detroit.
- McLaughlin, B. 1978. The monitor model: Some methodological considerations. *Language Learning* 28:309-32.
- Pease, D.M., J. Berko Gleason, and B.A. Pan. 1989. Gaining meaning: Semantic development. In: *The development of language*, ed. J. Berko Gleason. 2nd ed. Columbus, Ohio: Merrill.
- Popper, K. 1959. *The logic of scientific discovery*. New York: Basic Books.
- Reynolds, P. 1971. *A primer in theory construction*. Indianapolis: Bobbs-Merrill.
- Rice, M.L. 1983. Contemporary accounts of the cognition/ language relationship: Implications for speech-language clinicians. *Journal of Speech and Hearing Disorders* 48:347-59.
- Schouten, M. 1979. The missing data in second language learning research. *Interlanguage Studies Bulletin* 4: 3-14.
- Schumann, J. H. 1978. The acculturation model for second-language acquisition. In: *Second language acquisition and foreign language teaching*, ed. R. Gingras. Arlington, Va.: Center for Applied Linguistics.
- Spolsky, B. 1985. Formulating a theory of second language learning. *Studies in Second Language Acquisition* 7:3:269-88.
- Tannen, D., ed. 1988. *Linguistics and language in context: Connecting observation and understanding*. Norwood, N.J.: Ablex.
- Tannen, D. 1989. Talking voices: Repetition, dialogue, and imagery in conversational discourse. Cambridge: Cambridge University Press. To appear.
- Waldrop, M.M. 1989. An infinity of new universes. *Washington Post*, January 15, C3.
- Widdowson, H.G. 1988. Poetry and pedagogy. In: *Linguistics and language in context: connecting observation and understanding*, ed. D. Tannen. Norwood, N.J.: Ablex.

Less commonly taught languages: The current situation

Karin C. Ryding
Georgetown University

After five months of research on the topic of Less Commonly Taught Languages (LCTLs), I have come to the conclusion that problems in this particular field are both ill defined and underestimated. Because of the wide-ranging diversity of languages (the rubric 'Less Commonly Taught' applies to all languages except English, French, German, and Spanish), issues have either been viewed microscopically, as they apply to a particular language, or macroscopically, as they apply to the field as a whole. However, a great deal more remains to be done to identify needs, issues, and problems; to define subfields; to measure relative levels of linguistic difficulty and assess their pedagogical implications; to analyze commonalities and differences; and to indicate directions for further research.

In this paper I aim to identify some issues which, although tacitly understood by many, have not yet been the subject of explicit systematic analysis; to review some recent attempts to coordinate and systematize issues in the field; and finally, to indicate some directions for the future.

1 Identification of issues

1.1 Fragmentation. One salient problem that characterizes this field is that most professionals involved in the Less Commonly Taught Languages are focused primarily on issues pertinent to their own individual areas of pedagogy or research, which may be quite restricted. This is of course the case in many academic disciplines, but in other fields there is usually enough disciplinary coherence so that a central group is willing to undertake wider professional responsibilities such as editing and publishing journals, engaging in profession-wide research, and establishing and administering professional organizations. In the Less Commonly Taught Language grouping in particular (due to its nonexclusive nature), the diversity of interests, backgrounds, and disciplines has meant that the wider field has not yet emerged from a state which is both peripheral and amorphous.

Therefore, although organizations for Arabic teachers, Turkish teachers, Chinese teachers, Scandinavian and Slavic language teachers exist, there is no official or organized umbrella organization and little coordination among the various fields, many of which face common problems such as lack of funding for research, lack of resource materials, lack of trained cadres of teachers, low or irregular enrollments, difficulty of finding publishers for materials, and the

need for considerable administrative and public support for these disciplines in order for them to be visible and viable.

1.2 Fluctuating enrollments. LCTL enrollments are strongly affected by political and economic factors. Currently, Japanese and Chinese are enjoying an unprecedented wave of popularity, enrollments having shot up over 45% in Japanese and 28% in Chinese over the past three years.¹ Arabic enrollments increased substantially after the 1973 oil boycott, and remained at relatively high levels for several years. In the early and mid eighties, however, Arabic enrollments plunged as terrorism, extremist fundamentalism and hostage-taking reinforced negative stereotyping of Arabs and their culture.

Aside from the stress and disruptions that such irregularities cause in LCTL language programs, the short-sighted nature of public opinion should not be the only determining factor in enrollments. Long-range national policy goals need to focus on the essentiality of developing expertise in less popular but critically important languages in academic programs and in government training programs.

1.3 Career patterns of LCTL teachers. An issue which I have not yet encountered as a topic of formal discussion or analysis is the career pattern of LCTL teachers. Many of them engage in foreign language pedagogy without having academic degrees in their discipline, without having been trained in research techniques, literary analysis, linguistics, in methodology or in second language acquisition theory. I do not at all mean to imply that these professionals are less competent or less motivated than their academically trained counterparts. I know from experience that in many cases they are extraordinarily effective and dedicated language teachers. However, these professionals often do not teach as full faculty members in regular university programs, but in government and private language schools, whose needs and requirements are quite different from academic settings. Even at the university level, many are hired part-time or only on a temporary basis, often as an essential but minimal part of an area-studies program, and therefore have limited involvement with professional activity related to their fields. Moreover, many of them, discouraged by low salaries, irregular enrollment, lack of job security and limited career potential, do not 'network', nor do they find adequate opportunities to publish. They also tend not to join professional associations. The field at large is therefore missing out on the potential contributions of highly experienced language professionals, whereas they may be feeling that their lack of 'professional' status or academic credentials prevents them from gaining credit and recognition.

1.4 Identification of difficulty levels and analysis of LCTL-specific problems. One of the first steps toward defining pedagogical needs (such as length of training required for specific proficiency levels, psychological factors to be taken into account in teaching, and strategies for materials development) is an investigation and objective assessment of levels of difficulty presented by different languages or language families. Most LCTL teachers develop, after a certain amount of time, an accurate intuitive assessment of the number and type of difficulties in their own languages, as well as methods for dealing with

them. However, very few of us have openly discussed or researched these problems under controlled experimental conditions. An exhaustive analysis of these factors would require extensive research into the nature of the psychological and perceptual issues that confront the language learner and which are the source of varying degrees of frustration, blocking, and other negative factors which considerably detract from the learner's ability to attain communicative competence. Two issues which, as an Arabic teacher, I confront on a daily basis, are use of a nonroman script and the problem of diglossia.

1.4.1 Use of nonroman script. During approximately the past ten years, research on reading, in both first and foreign languages, has escalated dramatically. Yet there has been extraordinarily little research on reading issues for learners of nonroman writing systems, systems which present varying degrees of (im)penetrability to the language learner. Much more basic reading research is needed along the lines of an excellent article entitled 'Encoding Strategies Used by Native and Non-Native Readers of Chinese Mandarin,' by Edmund B. Hayes in the Summer 1988 issue of the *Modern Language Journal*, wherein the author revealed significant differences in native and nonnative reader processing of Chinese characters. For instance, he was able to pinpoint preferred strategies for native and nonnative readers, and also to raise the issue of the involvement of the right brain hemisphere in reading Chinese as a foreign language, with extremely interesting implications for further reading research.²

In the case of Arabic, for example, it was widely accepted for many years at the Foreign Service Institute that the script was so difficult that its introduction should be delayed, and the language taught initially in transcription. Moreover, students were not taught to write in Arabic, even though they eventually had to read Arabic script. This was justified by the task-oriented nature of FSI training, where writing in the foreign language is usually not a skill needed on the job, whereas reading is. Thus by the end of training, FSI students could understand an Arabic newspaper article, but couldn't write a simple sentence in Arabic script. Although this seemed to circumvent certain problems, it had the long-term effect of turning out students who were only semiliterate and unsure of graphic aspects of the writing system. Furthermore, although high-aptitude learners regularly achieved scores in the S-3/R-3 range after 40-80 weeks of training, very few students were ever able to achieve reading or speaking scores in the 4 range or above. While 4-level skills are not usually required for job performance, the fact that they are rarely achieved either as a result of in-country experience or advanced training means that the highest-ranking cadre of government-trained Arabists operates linguistically, with rare exceptions, below the advanced professional proficiency level.³

It has been my experience that although Arabic script is conceptually difficult to grasp to the point where it becomes second nature, teaching the language in transcription ultimately retards the acquisition of higher levels of communicative competence. Therefore, script should be used immediately, but under controlled circumstances and with specifically designed partial goals.⁴

1.4.2 Diglossia. Where the written and spoken languages are different, the standard and still widely accepted audiolingual emphasis on using spoken drills and exercises to reinforce reading doesn't apply. Nor can reading and writing skills be used to reinforce speaking and listening--at least not in the same direct and closely calibrated way. In the field of Arabic teaching, this has been the major stumbling block to developing effective proficiency-based materials geared toward communicative competence. Arabic teachers have yet to agree on what communicative competence is in Arabic, let alone reach a consensus on how to approach it. Charles Ferguson's 1971 article on 'Problems of Teaching Languages with Diglossia', in which he points out exactly these problems, is one of the few attempts to raise pedagogical consciousness and awareness of the cruciality of these issues to the progress of the field.⁵

1.4.3 Factor assessment. It should be possible to measure degrees of difficulty confronting English speakers acquiring a specific foreign language. For example, languages can at least be rated as to the quantity of conceptually unfamiliar issues which have to be faced by the learner, such as (1) nonroman script, (2) diglossia, (3) non-Indo-European origin, (4) complex inflectional system, (5) non-SVO word order, (6) unfamiliar phonological features. The presence of any of these features complicates the learning process, as has been known for a long time. They should now provide a basis for more empirical analysis, including precise measurement of the psychological and pedagogical impact of each factor. Such analysis should be a primary objective in LCTL research, and each factor could readily be the subject of more in-depth comparative and contrastive analysis. Only after issues such as these are thoroughly understood can they be systematically dealt with from the pedagogical point of view.

1.5 Current attempts at coordination and systematization

1.5.1 The GURT survey. In order to make this report as timely and relevant as possible, I developed a questionnaire on issues and problems in the LCTLs which I sent out to a number of colleagues in my own professional association, the American Association of Teachers of Arabic, and to colleagues in other LCTLs. Although I didn't even hope for half the questionnaires to be returned, the response rate was more like 20% than 50%, so I was not fortunate in my efforts to assemble a credible database. For the sake of initiating at least some empirical study, however, I would like to provide a summary of the responses that I received. (See Figure 1.)

Although these are only preliminary results, they show a strong concern with materials development, especially at the intermediate level of instruction, a difficult stage because of its transitional nature. Over 90% of respondents are influenced at least a little in their approaches to teaching by second language acquisition research, and overall communicative competence is the primary goal of most programs of instruction. Moreover, oral proficiency testing leads other types of testing. Taken together, these responses indicate

Figure 1. Summary of responses to LCTL questionnaire.**Question 1.** Most important issues in teaching your language.

Issue:	Average Score (1-5, 1=highest)
Materials	1.68
Teacher training	2.25
Low enrollment	3.2
Reference tools	3.36
Study abroad opportunities	3.55

Other issues mentioned: student attitude, use of visuals and computer, lack of support for students, lack of support from administration.

Question 2. Urgency of materials development.

Level:	Average Score (1-5, 1=highest)
Elementary	1.93
Intermediate	1.25
Advanced	2.31
Graduate	3.25 ⁶

Question 3. Extent of influence of second language acquisition research on your own teaching.

Extent:	Percent of response:
Greatly	38%
Somewhat	25%
Little	31%
Not at all	6%

Question 5. Teaching goals of your program.

Goal:	Average Score (1-5, 1=highest)
Overall communicative competence	1.38
Reading	1.6
Translation	4.0
Writing	3.33
Speaking and listening	1.73
Cultural understanding	2.69
Literary analysis	3.5

Question 6. Testing procedures in your program.

NOTE. All procedures received a substantial number of marks, indicating that many programs have mixed testing methods. In rank order they are as follows:

1. Oral proficiency
2. Reading proficiency
3. Discrete point/Translation [tie score]

a high level of professional involvement with second language acquisition research and its pedagogical applications.

2 Existing projects and resources. There are at least four projects which deal broadly with issues affecting LCTLs.

2.1 Center for Applied Linguistics (CAL): LCTL Data Bank. This resource consists largely but by no means exclusively of a continuing 'Survey of Materials' for Less Commonly Taught Languages, covering 1200 languages with 12,000 entries. These materials, as described by data bank administrator Dora Johnson, consist of 'basic tools of access' such as teaching texts, grammars, and bilingual dictionaries. Although much of the data is in a computer, it is not available on-line because of lack of funds to process the entries and convert them all to one system. The data collection for LCTLs began in 1967, when Charles Ferguson saw the need to make information available on what were then called the 'exotic' languages, and the first bibliography dealt with 'Study Aids in Critical Languages.'

Currently, in addition to the materials survey, CAL actively maintains an extensive journal collection, which is accessible at CAL headquarters for anyone interested in investigating particular LCTL issues. Moreover, CAL holdings include all the Peace Corps language training materials, which are not available anywhere else in the United States. The ERIC Clearinghouse provides either quick or extensive searches on specific subjects.

CAL is, of course, involved in many other activities that either directly or indirectly involve LCTLs, such as proficiency testing, language planning, country surveys, text development, work with refugees and minority children, and basic research.

2.2. LCTL Proficiency Guidelines. Developed jointly by ACTFL and the Center for Applied Linguistics, and published in 1987, this project was made possible through a grant from the U. S. Department of Education 'to motivate teachers within the LCTLs themselves to start defining and discussing proficiency-related issues and to draft the first provisional guidelines for several languages' (*Guidelines*, Preface). In addition to chapters on oral proficiency guidelines in general, there is an extensive bibliography about 'proficiency-related issues' and sections on proficiency measurement in Arabic, Hindi, Indonesian, and African languages.

2.3 National Foreign Language Center Conference. This annual conference has been held since 1987 at the National Foreign Language Center, currently headquartered at the School for Advanced International Studies, in Washington. This small group, with representatives from many national teachers' associations, meets annually to discuss issues directly related to LCTLs. This year, the meeting was held in February, and included panels on 'Empirical Research in Second Language Acquisition,' 'Technological Applications for Less Commonly Taught Languages,' and 'Design for Competency-Based Curriculum.' It aims at coordinating LCTL issues, sponsoring research and developing a national agenda.

2.4 The Stearns survey. In November, 1985, the State Department commissioned former Ambassador to Greece, Monteagle Stearns, to prepare a report focusing on proficiency in the four major 'hard' languages taught at the Foreign Service Institute: Arabic, Chinese, Japanese, and Russian. Ambassador Stearns interviewed more than 150 American and foreign experts, travelled to the areas where the languages are spoken, visited FSI's advanced language training schools in these countries, interviewed post language officers and discussed issues with principal officers and their deputies. He concludes that 'while we are doing well in giving a large number of Foreign Service personnel a smattering of hard languages, we are doing badly in developing real professional proficiency in those who most need it. Furthermore, I am convinced that our overall hard language capability is declining.' The Stearns report is a first-class, objective analysis of essential language issues that affect the conduct of foreign policy and the future of U.S. international relations. He makes 20 specific recommendations, including more in-country training, upgrading Language-Designated Positions (LDPs) to the 4 level at many posts, encouraging regional specialization for junior officers with previous hard language training, giving promotional preference to officers with hard language skills.

3 Directions for the future. Existing LCTL programs, including those which are 'currently decaying' (Pubillones 1989:6) need to be protected and extended, while long-term goals should aim at expanding and increasing awareness of the critical importance of LCTLs as well as developing a support structure for research, curriculum development, and training.

I believe there is a need and a market for an LCTL-specific professional journal to serve networking purposes as well as to publish current research and review relevant literature.

The sine qua non for all future activity in the field is funding. By this I mean not only fund-raising, but also encouraging LCTL professionals to seek grants, prepare grant proposals and submit them to appropriate funding agencies or foundations. Needs assessment, materials and curriculum development, proficiency testing procedures, and teacher training are all dependent on availability of funds, but most individual LCTL professional associations are ill equipped to monitor funding sources systematically, or to give professional guidance to their members who seek funding.

This means that there may be a need for a larger body, along the lines of a foundation, which can open itself to all LCTLs, and actively involve itself in funding issues. Other problems which can and need to be pursued are recruitment of teachers, sponsoring of publication projects, support of demonstration programs, introduction of LCTLs into high schools, international linkage programs (an overseas network is crucial to many LCTLs), collection of authentic materials, coordination and consultation with other organizations, lobbying, program and student evaluation, and also research into language learning anxiety, a topic which is particularly relevant for the more difficult languages.

Notes

I would like to express my appreciation to Joseph White, Rebecca Oxford, and Dora Johnson for their help and advice in preparing this paper.

1. Statistics taken from Collison (1988).
2. 'Perhaps, then, the right hemisphere is overloaded when Americans read strings of Chinese characters. The right hemisphere in American readers may not be as agile in processing the graphic patterns because this hemisphere is not used as much when reading an alphabet-based language like English. Moreover, an alphabet-based language provides far greater *graphic redundancy* (i.e., repetition of written symbols) than does the Chinese character system in terms of different graphic features presented to the reader' (Hayes 1988:193).
3. See the Stearns report for more detailed analysis of this problem.
4. See Ryding and Stowasser (forthcoming).
5. The term 'diglossia' was actually coined by Ferguson (1959).
6. Only half the respondents rated the graduate level at all, so its score is actually even higher than 3.25.

References

- Allen, Roger. Arabic proficiency guidelines. *Al-Arabiyya* 18.1/2:45-70.
- Argoff, H. David, Gary D. Crawford, and Hedy St. Denis. 1985. Oral language proficiency testing at the Foreign Service Institute. Unpublished paper.
- Collison, Michele N-K. 1988. Fascination with business and the Orient fuels enrollment in Asian languages. *The Chronicle of Higher Education*, February 17:A35.
- Daher, Nazih. 1987. Arabic sociolinguistics: State of the art. *Al-Arabiyya* 20.1/2:125-59.
- Ferguson, Charles A. 1959. Diglossia. *Word* 15:325-40.
- . 1971. Problems of teaching languages with diglossia. In: *Language structure and language use*, ed. Anwar Dil. Palo Alto: Stanford University Press.
- Foreign Service Institute, Department of State. (n.d.) Language skill level descriptions, reading and speaking.
- Hayes, Edmund B. 1988. Encoding strategies used by native and non-native readers of Chinese Mandarin. *MLJ* 72.2:188-95.
- Kubler, Cornelius C. 1985. Training for high level language skills. Draft.
- Lambert, Richard D. 1985. Foreign language instruction: A national agenda. *Foreign Language Annals* 18.5:379-383.
- McCarus, Ernest N. 1987. A Study of Arabic in the United States: A history of its development. *Al-Arabiyya* 20.1/2:13-27.
- Pubillones, Lillian. 1989. CAFLIS update: The dialogue continues. Coalition for the Advancement of Foreign Languages and International Studies. Report dated January 6. Washington, D.C.
- Ryding, Karin, and Barbara Stowasser. (forthcoming) Text development for content-based instruction in a non-Indo-European language. In: *Content-based instruction for foreign language classrooms: Methods and models*, ed. Betty Lou Leaver and Steven B. Stryker.
- Stansfield, Charles W., and Chip Harman. 1987. ACTFL proficiency guidelines for the less commonly taught languages. Washington, D.C.: Center for Applied Linguistics.
- Searns, Monteagle. 1986. Report on hard language proficiency in the Foreign Service: Arabic, Chinese, Japanese, Russian. Prepared for the Under Secretary for Management, U.S. Department of State.
- Swaffar, Janet. 1988. Readers, texts, and second languages: The interactive processes. *MLJ* 72.2:123-49.

Towards a rationale for language teaching technology

Peter Strevens

*Wolfson College, Cambridge
and the University of Illinois*

0 Introduction. Language teaching technology (LTT) is in widespread use. Expenditure on LTT has been very considerable, and growing, for some three decades, particularly in the more affluent countries. Yet in spite of the massive size of past and current investment in equipment for use in language teaching, few reliable figures exist about just how much has been spent by particular educational systems. There is still less in the literature that casts light on whether the investment is worthwhile, or that justifies (or refutes) an educational case for using LTT. This state of affairs is unfortunate enough in relatively affluent societies such as the United States and Great Britain; in developing countries it is doubly unwelcome, since (1) the absolute levels of expenditure represent a much higher proportion of total educational expenditure, and (2) the decision to make such expenditure is often taken by aid administrators in international agencies in the mistaken belief that language teaching equipment must, of itself and regardless of the competence of the teachers, lead to improvements in language learning.¹ The theme of this paper is that the cost-effectiveness of investment in LTT hardware is a function of the presence of (1) suitable software, and (2) teachers with appropriate expertise. In that respect it has affinities with some of the literature on language planning (e.g. Kennedy 1986) and appropriate technology (Markee 1986, 1989) and educational development generally (Beeby 1966). I shall add that on the one hand expenditure on LTT in developing countries is wasteful unless the teachers can make suitable use of it; but on the other hand the fact of being involved in software development, even at a low level of technology, can have a morale-raising effect on teachers, whose value is thereby much enhanced.

0.1 Two preliminary comments. Two preliminary comments are necessary. First, the paper does not address itself in detail to the question of how 'high-tech', 'low-tech'--or indeed 'no-tech'--the provision of LTT should be. The answer to that question is contingent upon particular circumstances, and especially (1) the level of educational development in the country concerned (roughly, upon economic affluence); and (1) the degree of professionalization of the teachers who will use the equipment.

The equipment itself exists on a scale of increasing complexity and sophistication, from the overhead projector and the photocopier at one end of the spectrum to computer-controlled interactive videodisc at the other. If

the teachers are to make effective use of the equipment, their professional understanding must match this technological sophistication. This requirement is essential, because LTT, in order to be effective, has to be teacher-driven: it cannot operate without the guidance of skilled teachers.

Second, the paper rests upon a set of assumptions about the nature of language teaching and the task of the teacher. One assumption is that all non-impaired human beings can achieve a reasonable practical command of any foreign language. Not a 'native-like' command, be it noted: this high level of competence is in fact increasingly common in 'second language' conditions, such as when English is learned and greatly used in Singapore or Hong Kong, or French in the Ivory Coast or Algeria. In second language conditions a native-like command of the language is reached by relatively large numbers of adults, no doubt as a consequence of social and sociolinguistic conditions subsequent to school learning. But as far as 'foreign language' conditions are concerned, to the extent that language teachers believe that achieving a native-like command of the foreign language is a realistic goal, they hold unreasonable expectations. Only a tiny proportion of learners even wish to achieve a native-like command; yet all learners can achieve more modest goals, which are still worth attaining for practical purposes.

Another assumption is that a basic task of language teaching is to enable the language learner to attain his/her 'maximum rate of learning', for as long as possible. Next, the assumption that there are many ways in which a professional teacher can deliberately promote an increase in the learner's rate of learning, and that one of these ways is by the introduction and effective exploitation of language teaching technology. The final assumption is that LTT serves only as 'an adjunct to the teacher': the effectiveness of the teaching depends on the teacher, not on the technology. Only well-prepared teachers, using suitable software, can make effective use of language teaching equipment. The equipment by itself is of virtually no value unless it is employed by a teacher who knows how to use it.

0.2 Divisions of the paper. The paper divides into three parts. First, it suggests that cost-benefit analyses in this field of education are feasible, and that they can give an indication as to whether large-scale expenditure on LTT can be justified in terms of results achieved; this section concludes with a case study of the costs of LTT in one large British EFL organization. Next, the paper offers a reappraisal of the role of LTT within the LL/LT process, and a 'best-case' assessment of the cost-effectiveness of LTT; it concludes that for an increase of 10% in teaching costs the potential yield of LTT is a net overall gain in the quality and level of learning achievement, of the order of one-third. However, the paper also cautions against over-investment where the economic and professional conditions are not appropriate.

1 Cost-benefit analyses of LTT

1.1 What is covered by the term LTT? Within the definition of 'language teaching technology' at least the following are typically included (see Table 1).

(1) **Sound reproduction and recording equipment.** In practice, the great majority of expenditure under this heading goes on audio cassette recording equipment, and much of it is of 'speech quality', not 'broadcast quality'. That is, LTT includes quite low-level electronics as well as sophisticated equipment.

(2) **Listening centers.** This includes sets of replay-only audio cassette machines, backed by collections of specialized recordings in the foreign language and by student worksheets, exercises, assignments, etc. These are in effect libraries of recordings, accessible to the learners, supported by teaching materials, and largely devised and written by teachers in the same institution: listening materials are not yet commercially available on a large scale, although the extent of such materials is growing.

(3) **Language labs.** There is a fashionable view of language labs that regards them as (a) inherently behaviorist and (b) pedagogically ineffective. But to believe this is to ascribe to mere hardware the philosophical stance of the writer of materials for use with language labs; it is also to misunderstand the potentialities of language labs, which are as effective as the materials used in them. Language labs are best seen simply as 'specially equipped classrooms'. If the tapes used are varied, interesting and relevant, good results can be achieved.

(4) **Video recording and replay equipment.** This includes equipment for use along the spectrum from simply replaying commercially made video films for teaching, through teachers making their own videos, to teachers assisting the learners to make the videos as part of 'experiential learning'.

(5) **Micro computers.** Teenage and young adult learners are already familiar (in many countries) with computer games and even with personal computers. Exploiting this knowledge as an adjunct to language learning is proving effective. Most software has to be specially written; many teachers also claim good results from teaching the skills of word processing in the foreign language.

(6) **Videodisc.** In the near future it will be necessary to add videodisc to this list of LTT, particularly in the form of interactive, computer-controlled videodisc.

(7) **Other LTT items.** The electronic and electro-mechanical equipment listed above is typically seen as the extent of language teaching technology. But a new group of LTT items have recently made themselves indispensable. At least in European EFL and foreign language teaching, teachers now commonly supplement and supplant their conventional coursebooks with materials of their own devising. Hence simple graphics materials, overhead projectors and transparencies, and photocopying facilities are seen as essential services to be available to the teacher.

Table 1. Categories of LT technology.**Principal types of language teaching technology:**

- 1. Overhead projectors and transparencies; photocopiers.** The basic tools that enable the teacher to supplement, and eventually to transcend, the limitations of the coursebook.
- 2. Sound reproduction and recording equipment.** Principally using audio cassette recorders, but also as elements of more complex items, such as video.
- 3. Listening centers.** Collections of replay-only cassette machines supported by a growing bank of recordings, of the widest possible range, usually backed by worksheets and other adjuncts to focus the students' learning.
- 4. Language labs, with accompanying cassettes and printed materials.** Nowadays the best of these materials are often prepared by the teachers, not purchased on the commercial market.
- 5. Video playback and recording equipment.** With accompanying facilities for editing and titling.
- 6. Microcomputers.** With a range of software, from video games through specially written programs to teach particular language points, to word processing.
- 7. Computer-controlled interactive videodisc.** Not yet widely available, but the subject of much active research.

1.2 Justification for LTT expenditure. Major expenditure by a language-teaching organization for any purpose should be based on a cost-benefit analysis. In the case of LTT--and indeed throughout much of educational expenditure--such analyses are hardly ever attempted. It should be realized that any big investment implies a tacit planning decision; it assumes that an expenditure of x thousand pounds will eventually produce results whose benefits to the organization (by way of e.g. improved learning, increased student satisfaction, hence higher recruitment and greater tuition income) will be no less than x thousand pounds, and probably will be greater than that. In short, it is taken on trust that the value of the benefits will exceed the costs.

In what follows I shall attempt a cost-benefit analysis of the investment in LTT by one major organization in the field of English language teaching: the Bell Educational Trust, of Cambridge, England. First, I shall briefly outline how costs and benefits may be assessed.²

1.3 Costs.

(1) Capital costs. The raw 'capital costs' of LTT equipment are not of themselves very illuminating. They become more interesting if it is remembered that they relate to 'teaching costs' (as distinct from e.g. property costs, travel and transportation, provision of meals or other student services, etc). Further, LTT costs include not only 'capital' expenditure but also a great deal of supporting costs, notably: spares and software, and directly related staffing and labor costs (e.g. technicians, teachers who may be engaged

full-time on LTT work, staff release time for developing teaching materials, etc). Hence in expressing LTT costs it is necessary to sum all these major categories of LTT expenditure.

Table 2. Categories of probable LTT costs.

Capital costs:

Costs of equipment purchase, with depreciation allowances where appropriate

Support costs:

- (1) maintenance contracts, spare parts, etc.
- (2) blank tapes, software
- (3) technical and maintenance labor
- (4) time allowances for teachers with respect to writing teaching materials, familiarization with techniques, trials of new materials, etc.

(2) Capital costs in relation to teaching costs. Summed in this way, LTT costs can be related to the other categories of teaching costs, for example, to the average annual gross costs of the teaching force. In most educational organizations a crucial statistic is the number of income-generating students and the number of teaching staff they justify. Hence it will be particularly insightful to relate LTT costs to the unit cost of a single teacher or full-time equivalent. Expenditure on LTT equipment and its support can then be expressed as equivalent in cost to a given number of teachers. Thus a broad index of the cost of LTT to an organization can be furnished by the ratio of: (a) total costs of LTT purchase and support, to (b) gross teaching costs. The relationship can also be expressed in the form: 'Every *n* teachers are accompanied by 1 full-time teacher equivalent in LTT costs'.

1.4 Benefits. The benefits to an organization of possessing and using LTT are not easily quantifiable. They depend on concepts such as: improvements in learning performance, whether through reaching the same target in a shorter time, or reaching a higher target in the same time, or embodying some aspect of improved effectiveness of learning (e.g. greater certainty of grasp, quicker comprehension, increased resistance to attrition, etc.); greater student satisfaction leading to improved reputation and increased future recruitment and therefore increased tuition revenue; improved teacher morale through pride in achievement, being professional and market leaders, etc; and similar rather nebulous quantities.

One aspect of the outcomes of investment in LTT, namely, observed improvements in learning effectiveness, can be isolated. Highly experienced teachers agree broadly on the levels of achievement that are likely to be displayed by various types of student after a given period of learning, without LTT and following its use. In short, they recognize a typical 'learning yield per unit of teaching time', and can estimate parity, shortfall, or improvement with

respect to the pre-LTT norm. In addition, experienced teachers are in a position to assess with reasonable accuracy whether the quality of language achievement post-LTT contains any elements that would not have been present without LTT.

Two further aspects of the benefits accruing from the use of LTT can be observed. First, it is noticeable that for many students some categories of LTT (notably listening centers, video and computers) are inherently of so much interest, and exercise such a strong attraction, that they feel impelled to voluntarily devote a significant number of additional hours to the learning task, that is to say hours of their own time, in the evenings, or at weekends. This addition to the previous norm of learning time can become a valuable and regular increase in total instruction time, and hence is an increase in both 'total quantity' and 'intensity of instruction'. Second, the photocopier and the overhead projector between them have revolutionized the nature of what goes on in language teaching classrooms. They have opened the doors of the classrooms to 'authentic' materials and have liberated teachers from the shackles of the coursebook. They have helped teachers to realize that any good teacher can improve upon any particular segment of any coursebook. This awareness has been a great stimulus to teachers' imagination, creativity and professionalism.

Table 3. Categories of potential benefits from the use of language teaching technology. (Full range achievable in the 'best case' only.)

Potential benefits:

- (1) Increased quantity and/or intensity of instruction
 - (2) Instruction transcends teachers' personal abilities
 - (3) Accelerates professionalism of teachers
 - (4) Increased learning yield per unit of teaching time
 - (5) Improved exit performance of learners
 - (6) Increased student satisfaction
 - (7) Improved teacher morale and hence effectiveness
 - (8) Enhanced reputation for the institution
 - (9) Improved recruitment and marketing prospects
-

1.5 A case study of capital outlay and recurrent support costs: LTT in the Bell Educational Trust. Now let us consider a case study, based on a particular large-scale institution engaged in the teaching of English as a foreign language.

The Bell Educational Trust (BET) is a nonprofit educational charity based in Cambridge, Great Britain, and specializing in the teaching of English as a foreign language. Its figures for expenditure on LTT over an eight-year period through 1988 are given in Table 4. BET's six schools of EFL in Britain teach some 42,000 student/weeks per year. In addition, BET has teachers on several language training projects in developing countries. The Trust has the reputation of being a leader in EFL teaching, both in methodology and in teacher training. The level of professional training required in its teaching

staff is second to none, and it probably makes the highest level of LTT investment to be found in a British--or indeed any other --EFL organization. This implies that a technical judgment has been made by first-class EFL professionals in a successful, large-scale operation, that expenditure on LTT is worth its cost. But while the expenditure has indeed been deliberately planned, the rational basis for making it has been at best sketchy. The chief justification for spending large sums on LTT has been the enthusiasm of excellent EFL teachers to improve pedagogical results through (among other things) technological innovations in teaching, followed by a vague, unarticulated judgment that the results constitute an improvement in learning.

Table 4. Annual costs of purchasing and supporting language teaching technology in the schools of the Bell Educational Trust, Cambridge, 1980-1988.

Elements of annual recurrent costs of LTT:

	£	\$
Capital: averaged over 8 years:	25,000	45,000
Support costs:		
spares and software	28,000	50,400
staffing and labor	<u>122,000</u>	<u>220,000</u>
TOTAL:	£ 175,000	\$ 315,000

Comparison:

This annual total is broadly equivalent to:

- (1) the gross cost of 11-12 tenured teachers, i.e. some 10% of the teaching force
 - (2) 7% of all teaching expenditure
-

The LTT equipment currently in use in BET schools covers the full range outlined above. Every classroom has an overhead projector; facilities for making transparencies are everywhere available; photocopying is constantly at hand. Each school possesses at least one, usually two, language labs, typically of 16 places, together with a listening center; all schools have a video studio with several cameras and editing facilities; all schools have a number of micro computers and a large collection of EFL software. It is worth noting that the machines used are the BBC B or BBC Master computers, manufactured by Acorn Computers. Almost all British EFL computer work up to 1989 has been standardized on these machines. However, as the BBC B machines become obsolescent and as multilanguage programs become available, other makes of computer are likely to be more widely introduced in future. Several members of staff are engaged in the development of computer-controlled interactive video materials on a collaborative basis with the British Council and other organizations.

Over a period of eight years to the end of 1988, the Bell Trust invested approximately £200,000 (\$360,000) in LTT capital equipment, i.e. an average annual rate of £25,000 (\$45,000).

Annual support costs average £28,000 (\$50,400) for spares and software, with staffing and labor costing a further £122,000 (\$220,000). Total annual costs for LTT are thus £175,000 (\$315,000).

This is equivalent to the gross cost of 11-12 tenured EFL teachers, nearly one-tenth of the total permanent teaching force of the Trust. So it can be stated that broadly for every 8 or 9 tenured teachers the Trust spends on LTT the cost of one extra teacher. This is a heavy financial commitment to LTT. The question is whether the expenditure is justified.

2 Functions of LTT in the LL/LT process. In the most general terms, LTT constitutes a part of educational innovation, an element in the professionalization of teaching, and a means to improve both the rate and the quality of the learning process. In more detail, the functions of LTT today can be seen to include the following.

- The simplest forms of LTT (OHP and photocopier) have become the typical enabling conditions for class teachers to transcend the traditional iron grip of the coursebook. Given these tools for doing so, teachers quickly discover that they can improve on the conventional coursebook, and thereby can extend their creative imagination as classroom practitioners.
- LTT of the more electro-mechanical kinds is an adjunct to the class teacher, extending his/her possibilities of performance in the foreign language -- e.g. by making it possible to present the learners with different voices and other varieties of the language, access to recordings (both audio and video) of great role models and performances, a wider range of contexts through the great potentiality of film and recorded documentary and drama. It can also compensate for possible weakness in the teacher's command of the FL.
- As for the learner, LTT provides means of offering to the student additional ways of communicating and expressing him/herself in the foreign language beyond the conventional activities of response to the teacher, of written or spoken exercises, of essays and compositions, etc.
- LTT keys into the nexus of organizing concepts which the profession has developed in recent decades (approach, methodology, syllabus, materials, evaluation, teacher training), chiefly at the levels of syllabus, methodology and teaching materials. This integration of technology into the organization of LT is a consequence of, and contributes to, the growth of professionalism in teaching.
- In terms of teaching input, LTT offers possibilities of greater scope, authenticity, contextualization, consistency, range of sources, relevance and significance to the learner. In addition, and particularly when visual input is added to auditory, LTT permits the teacher to present to the learner (or enables him to obtain it for him/herself) much of the 'schematic' information (information about 'the nature of the world') which provides the necessary contextual meaning to the systemic information (about language). This is a powerful addition to the usual techniques of teacher plus coursebook alone.
- In relation to the learner's 'intention to learn' (Strevens 1988), LTT gives powerful new means to the teacher, and to the self-motivated student, to maintain his/her interest in learning and to keep up the on-line mental effort necessary in order to progress as a learner. It does so by offering opportunities for the teaching material to be invested with those qualities--

interest, variety, humor, challenge, impact--which typically convert a reluctant learner into a willing one.

- LTT greatly extends the teacher's opportunities for practice and use of the language. The performance skills of language, like all human psychomotor performance, require frequent occasions for rehearsal, practice and use. LTT makes possible performance training activities, in the skills of reading, writing and speaking, that cannot otherwise be contrived. This is particularly true in regard to the chance of exact repetition, the making of recordings of the student's own efforts for later comparison, etc. The use of sound recordings has often been thought of as affecting only phonetic and phonological control by the learner. But whether that is so depends crucially on the focus of the learner's attention: if he/she is concentrating on listening to phonetic performance, that is what will be keyed into the learning processes; if on lexis, or on grammar, or on aspects of discourse, then they, too, will key into the learning. The suggestions and instructions given to the learner when starting a session of use of recordings will strongly affect the linguistic area within which learning takes place. Listening centers, language labs, video, computers etc. increase the range of interesting possibilities for this essential part of the language teaching/learning process.

3 Is LTT worth the cost? In one sense, the answer must be a subjective one. Yet in another sense it is possible to make a professional assessment of some of the qualities already noted. The Bell figures, it must be stressed, represent a 'best case', since it is most unlikely that more than a handful of organizations have made an equivalent investment or achieve higher learning standards. Indeed, it is important to emphasize two points: (1) it is not being suggested that other language teaching institutions should necessarily seek to emulate the Bell level of investment, unless such a level is appropriate in their circumstances; (2) such a high level of investment is only indicated if the standard of professionalism among the teachers is also of the highest. Buying expensive equipment without the software to use with it, and without teachers of the highest calibre, is likely to be pedagogically valueless.

In this 'best case', then, LTT represents an increase of broadly one-tenth in total teaching costs. The crucial question is whether the value of the learning benefits discerned reaches this level. If not, LTT is a net loss; if it exceeds it, LTT represents a net gain, at least in this 'best case'.

The benefits can be estimated as follows:

- extra learning time, from heightened interest: 25%
- improvement in quality of learning not available from teacher-only instruction: say 15%
- overall improvement in learning achievement, taking into account general confidence and breadth of command of the language: 35-40%
- heightened 'customer satisfaction' reflected in lack of complaints, increase in numbers of students applying as a result of personal word-of-mouth recommendation, and hence improved economic performance: 10%

These percentages cannot, of course, be added to each other since they are percentages of different quantities. But from the point of view of the learner--the fee-paying student or his/her sponsor--one may reasonably ask what proportion of the value of the learning experience (i.e. of the tuition fee and other costs he/she has paid) can be attributed to the presence and utilization of LTT? The question goes to the heart of the quality of the learning experience, and to the value perceived by the student for his outlay, and it seeks to identify a proportion of that value as dependent upon the use of LTT. At a conservative assessment this would be at least 30%.

If this figure is reasonably close to being accurate, it suggests that the cost-effectiveness of language teaching technology, at least in the best case, is at a satisfactory level, and well above its net costs. As so often in the large-scale organization of language teaching, it transpires that 'there are some economies the poor cannot afford': if the teaching force is of the highest calibre, then it is financially worthwhile making a big investment in LT technology; if the teachers and other indicators are not above average, investment in LTT is best kept to a lower level and allowed to rise in step with an increase in general professional standards. In short, in the best case it can be maintained that an increase of 10% in teaching costs can result in an added value to the learning product of some 35%.

Notes

I am greatly indebted to my former colleagues, Keith Morrow, Director (Education), and Pat Turner, Director (Finance) of the Bell Educational Trust, for assistance in preparing this paper. It is they who are largely responsible, through the management of pedagogical innovation and efficient financial administration respectively, for the considerable success achieved in the Trust's schools as a result of the widespread introduction of language learning technology. I am grateful also to Dr. N. Markee for helpful comment and discussion, particularly in relation to LTT in the conditions of 'less developed countries'.

1. Since the 1950s the provision of every kind of hardware remotely connectable with language teaching has been part of international aid to developing countries. This is perhaps not surprising, since equipment represents to an aid administrator--who is not typically a language specialist--a single, once only purchase cost. Donors prefer such items rather than paying the open-ended recurrent costs of people, such as specialist teachers, for example. The Third World retains sardonic memories of language labs exported unheedingly in the 1960s to places where the professional understanding to use them was lacking, where there was a paucity of suitable teaching tapes for use with them, where there were no technicians to service and maintain them, and even, in a few cases, where there was no electricity supply to run them. Similar waste still occurs today in the unthinking supply of video equipment and computers (for language teaching) to countries before the conditions exist for making proper use of them.

See Kennedy (1986) and Thorburn (1971) for discussion of investment in language planning, Markee (1986) for discussion of 'appropriate technology', and Beeby (1966) for an outline of stages of educational development. The

issue is not, of course, that LTT should never be supplied as part of educational aid, rather that aid administrators should recognize that it requires a careful judgment on the part of professional language teaching specialists to determine whether the circumstances exist that would make a big investment appropriate. The technology, of itself, does no teaching. The teachers, if adequately trained and experienced, can improve their effectiveness by means of LTT.

2. It should be noted that Chris Kennedy's admirable 1986 paper, 'Costs, benefits and motivation in ESP', uses the term *cost-benefit analysis* in a metaphorical sense rather than a financial one.

References

- Beeby, Charles. 1966. The quality of education in developing countries. Cambridge, Mass.: Harvard University Press.
- Kennedy, Chris. 1986. Costs, benefits and motivation in ESP. Mimeo. Paper delivered at the University of Algiers, March 1986.
- Markee, N. 1986. Toward an appropriate technology model of communicative course design. English for Specific Purposes 5.6:161-72.
- Markee, N. 1989. The notion of vulnerability: A key concept in innovative program design. Mimeo. Urbana: University of Illinois, Division of English as an International Language.
- Strevens, P. 1988. Learning English better through more effective teaching: Six postulates for a model of language learning/teaching. World Englishes 7.1:51-61.
- Swales, J. 1986. Service English program design and opportunity cost. In: Program design and evaluation, ed. K. Johnson. New York: Prentice-Hall.
- Thorburn, T. 1971. Cost-benefit analyses in language planning. In: Can language be planned? ed. J. Rubin and B. Jernudd. Honolulu: East-West Center Press.
- Tollefson, J.W. 1981. The role of language planning in second language acquisition. Language Learning 31.2:337-48.

ESL program evaluation: Realities and perspectives

Ali Hajjaj and Balkees Al-Najjar

Kuwait University

This paper discusses some issues that highlight the importance of systematic ESL program evaluation. Guided by these issues and an investigation of ESL evaluation in higher education in the Arab world, a set of perspectives (after Miller 1988), related to plan, process, data, action orientation, and evaluation of evaluation, is suggested.

1 Introduction. Program evaluation constitutes no problem in its general sense. In its specialized sense, it poses questions concerning its nature, purpose, scope, practices, and products. Is program evaluation a formal assessment of the worth of educational phenomena, the determination of how objectives are realized, or the provision of information for decision making? (Cf. Popham 1975, Tyler 1950, Stufflebeam et al. 1971, respectively.) Evaluation may imply all these definitions, but it also means, for example in relation to ESL programs, what Cronbach et al. suggest, 'a systematic examination of events occurring in and consequent of a contemporary program--an examination conducted to assist in improving this and other programs having the same general purpose' (1980:14).

There is also a considerable amount of technical terms used, concepts formulated, and a 'galaxy of activities' pursued (Harlen 1974:128). For example, the scope is macro- or micro-; the phase is initial, formative, summative, or longitudinal; data-processing is either descriptive, comparative, or analytic; the status is either responsive, illuminative, or transactional; and so on (cf. Al Sani et al. 1981, Popham 1975, Worthen 1985, Nevo 1985). Although these 'artificial' divisions may reflect different evaluation attitudes or approaches, from the point of view of ESL programs, they must not be overstressed. 'They may all be relevant...[but] it is essential that any type which is relevant is used at the appropriate time' (Harlen 1974:129).

2 ESL program evaluation

2.1 The present situation. ESL program evaluation is a newly emerging aspect of the general evaluation field (cf. Saltzer 1982:89, White 1988:148). Within language curriculum development, evaluation is beginning to receive greater attention. The following profile seems, in our view, to characterize the present state of ESL program evaluation.

First, whereas general terms and concepts are widely used (cf. Perkins and Angelis 1985), new ones, specific to ESL programs, are getting into the

literature. For example, a distinction is now drawn between process and product evaluation rather than between formative and summative evaluation (but cf. Popham 1975:33). Long (1984) suggests that process evaluation can address issues that are often ignored by formative studies, such as the procedures and findings of classroom-centered research on language learning. Berretta (1986:441) makes a case for what is called program-fair evaluation. He argues that '...the claims of each specific program must be taken into account in test construction if competing interests are going to be presented.' On the other hand, a communicative curriculum would, according to Breen and Candlin, rely on 'shared and negotiated evaluation' (1980:104).

Second, tests and examinations have long been employed as synonymous with evaluation, and, in many situations, as the only criterion for measuring program effectiveness. However, they are increasingly seen as only one aspect of evaluation (cf. Yalden 1983:96, Al Busairi 1988). Many affective and cognitive aspects of learning cannot easily be assessed by examinations, e.g. oral work. Such parts should be assessed because otherwise they might not be taught (cf. Dines 1984:122).

Furthermore, although much emphasis is placed on tests as an evaluation of students' performance and, by implication, the curriculum, test results are not the only criteria for judging a program. According to Becher et al. (1980), the main limitation of testing is its 'inability to throw any light on causes and effects. Test results assert, but never explain' (quoted in Holt 1981:150).

Third, there is a shift from the concept of evaluation as a final phase of the process of developing language programs (Yalden 1983:96) into that of seeing it as an integral part of curriculum development in its various stages and aspects (cf. White 1988:148). Within the communicative language teaching methodology, evaluation is no longer a summative, end-of-course activity, but rather 'an immediate and ongoing activity inherent in communication and in learning to communicate... genuinely formative for the individual learner and the whole curriculum' (Breen and Candlin 1980:105). In this respect, the evaluator becomes an integral part of, rather than a specialist somewhat outside, a project.

Finally, since ESL program evaluation, like other types of evaluation, entails making judgments, ethical questions such as standards, justification, value and utility, both in evaluation and evaluation research, become most essential (cf. White 1988).

2.2 The need for ESL evaluation. As a preparation for investigating ESL program evaluation in higher education in the Arab world, the need for such evaluation will now be exemplified with reference to Kuwait University (KU).

2.2.1 Placement tests. Students enter KU on the strength of their overall percental achievement in all high school subjects, including English. English placement tests, based on college English courses, are then administered. The students who achieve a certain percentage are placed in credit courses and the remaining students go into remedial studies. The question is: is it fair for the students to be placed by a test whose content is what they have not yet studied? Second, the test consists of three multiple choice parts--grammar,

vocabulary, and reading--only, thus ignoring students' listening and speaking abilities. Third, high school English language textbooks claim to be written with a communicative rather than a grammatical approach in mind--a fact also disregarded by college placement tests.

An attempt is currently being made to standardize a test based mainly on English taught in preuniversity education. But even in its present experimental stage, it is clear that this new test is much less difficult for certain categories of students (e.g. those majoring in engineering) than college placement tests. This implies that the new test, once operational, will allow low-level English students to enter into credit courses. Such a situation reflects indecision concerning which content to use for placement tests: high school English content or college English content. A formal/systematic evaluation study conducted on both types of English and both ESL situations could provide a fair solution (cf. Safi 1986, Berretta 1986).

2.2.2 Classroom interaction. There are indications that students are not learning English to the level required: they are unable to make sufficient use of classroom teaching in their specialization and they find it difficult to use their knowledge of English in their project work, or later in job situations (cf. KU Language Center's Users' Report, 1988). Part of the reason may be that we do not know much about what goes on in the classroom, what students are learning, or what language or languages are being used. We do not know how effectively feed-back on students' learning is used by teachers to identify problems and suggest solutions. All this is not captured by end-of-term tests, for example. But it can usefully be captured by continuous assessment illuminated by the findings of process evaluation (cf. Long 1984).

In this respect, there are several other issues which necessitate conducting ESL evaluation: the use of English as a medium of instruction in science, engineering, and medical studies as matched against the use of Arabic in social studies; the excessive heterogeneity of the proficiency levels of students (cf. Gaffney and Mason 1983); and the role of specialization teachers in developing students' English, among other issues.

2.3 A closer look. We now come to an investigation that aimed at identifying ESL evaluation realities and future prospects in KU and a number of other Arab universities. A questionnaire was addressed to language center directors, heads of departments, language unit supervisors, and teachers. Responses have shown that the existing evaluation pattern is unsatisfactory, but there is an ambition not only to have evaluation as an integral part of the ESL process but also to have its findings more effectively utilized. The following are the specific findings and conclusions of the investigation.

2.3.1 Specific findings. Only in a few cases were there no evaluation studies. Where such studies were conducted, they were mostly concerned with program modification but much less with establishing new programs. The purposes of evaluation were: (1) to confirm the validity of objectives, methods, and materials used, (2) to establish the real needs of students, and (3) to develop programs to meet specific student needs.

Current evaluation focuses mainly on language content; to some extent, on needs analysis, teaching techniques and procedures, and students' achievement; and to a lesser extent on testing techniques and procedures, students' in-course performance, instructors' competence, teaching media, student-teacher relations, and subject-teachers' contribution to the improvement of students' English. Some important aspects such as learning strategies, library-/computer-related activities, students' on-job English performance, and, most importantly, administrators' competence have received little or no attention. In a future evaluation, all the above-mentioned aspects are considered, by most respondents, as necessary components. Some respondents, however, specified program effectiveness, proficiency of new students, test design, and item, part, and whole test effectiveness as central for the evaluation project.

As for evaluators, the present tendency of involving internal and external experts is also acceptable for the future. Yet there is a strong desire among the respondents that evaluation should mainly take the form of institutional self-study in which all program parties are involved.

In terms of data collection, the commonly used techniques are: analysis of textbooks, instructors' oral/written reports, students' answer papers and projects, and questionnaires. Interviews and audio/video recording of classes are among the techniques suggested for future evaluation. Some respondents added that evaluators' common sense or intuition should also be used (cf. Miller 1988).

Evaluation conducted so far was described by respondents as formative, summative, or a combination of both types, with the latter being favored for the future. Evaluation results have been used, in most cases, in adapting existing programs, in preparing new materials, or in revising objectives, teaching techniques, and tests. In some cases, evaluation results were not used either because of their insignificance, the limited scope of evaluation, or the lack of interest in evaluation and its results. Most of the respondents, however, agree that evaluation should be conducted regularly and that teachers should be trained for such a task.

Finally, only in less than half of the cases investigated was there an evaluation of evaluation. In the future, most of the respondents want to maintain or introduce this activity with internal evaluators, and, to a great extent, external experts to be involved in it.

2.3.2 Conclusions. The findings mentioned above reflect a positive attitude towards evaluation and a desire to establish it as an integral part of ESL programs. These findings also reflect an awareness that an evaluation component may effectively improve program processes and students' English proficiency. If such is the case, why has not ESL program evaluation received its due attention?

The following points may provide an answer.

- (1) The meaning of evaluation is often associated only with end-of-course examination results or with program modification.
- (2) Involvement in evaluation is often limited to certain parties (e.g. course-teachers and language unit supervisors). What is required is the

involvement of all parties concerned: students, teachers, supervisors, administrators, as well as specialists.

(3) There are certain variables which are rarely considered such as supporting ESL programs with adequate budgets, the status of instructors and the material incentives to sustain their motivation and interest, and the frequent turn-over of ESL instructors in Arab universities which hinders the smoother continuity of programs. Ignoring these and similar variables makes program evaluation a difficult task.

3 Perspectives. There are many indications that favorable conditions exist for further development, research, and evaluation in the ESL field in the Arab world. English continues to play an increasing role, and demand for it for educational purposes, economic and social activities, and international understanding, is growing. There is also a persistent desire to improve on the outcomes of English language education at all levels: in preuniversity, university, and on-job settings. Systematic ESL program evaluation, as yet to be comprehensively attempted, can effectively contribute to such improvement. Consistent with this orientation is what Davis (1980) calls 'the increasing exploration of methods, materials and environments for improving and broadening student learning' (p. 129). In this respect, KU Program Academic Evaluation (PAE) Project (1988) warns that the damage that may be done to the nation in the case of the deterioration of its higher education cannot be compensated.

The following perspectives (based on Miller 1988, but also cf. KU PAE Project 1988) are seen as a useful framework to suggest for future ESL program evaluation.

3.1 An overall evaluation plan should be developed and communicated. Central to students' English language proficiency is the question of testing. Therefore, first and foremost, all possible means should be taken to improve upon the quality of the tests given. In an ESL program, various tests should be involved: high school leaving tests, university entry placement tests, common achievement tests taken by sections of the same course, and quizzes meant to monitor students' growing knowledge and skills.

Tests can provide an immense feedback on how well students perform and how well courses function. However, tests are not the only source of feedback. An overall evaluation plan should seek feedback from other components as well as from all persons involved in a program. Another reason why such a plan should imply more than testing is the fact, referred to earlier, that there is a great deal of dissatisfaction with students' achievement in high school and university as well as in the job market. That dissatisfaction cannot be blamed only on students, teachers or administration since sufficient evidence on where weaknesses lie is lacking. A comprehensive evaluation plan, therefore, should include: (1) language policy: leadership and budget; (2) administration: personnel needs, administrative efficiency, and relations; (3) teaching: methodology, procedures and techniques, training and development; (4) students: testing, attitudes, abilities, performance, and needs; (5) curriculum: required courses relevant to objectives, content, facilities, and media.

One important aspect of such a plan is that students, teachers, supporting staff, and administrators should realize that they all contribute to the evaluation scheme in its various stages, and benefit from its outcomes. For example, as far as students are concerned, there is a need 'for an instructional program that deemphasizes students' total reliance on memorization and examinations for completing course requirements' (Safi 1986:421). Such a program is likely to be achieved when the processes involved are as important as their outcome.

3.2 The process of evaluation is as important as the product. According to KU PAE Project (1988), evaluation, as a mental and intellectual process, is an end in itself. It provides educational institutions with experiences that can only be acquired through evaluation. Such experiences cannot effectively be gained if evaluation is limited to the end-results of programs. Rather, it should be concerned with monitoring progress throughout the program. Then, feedback becomes a built-in and cumulative aspect which will, in turn, establish evaluation as an integral part of the continuously pursued activities of the program. In this way, evaluation becomes concerned with 'the changes the program brings about in teachers' and students' attitudes, students' self-concepts, related intellectual skills, and the like' (Long 1984:409).

Unlike academic content programs whose primary object is the subject matter, in ESL lessons, 'language is both the vehicle and the object of instruction' (*ibid*). Process evaluation provides information on language learning strategies, classroom interaction, affective aspects of learning and the type of language used, over and above the final view on courses obtained through product/summative evaluation. In this respect, an illustration is in order: Branson et al. (1986:70) claim that it has been possible to improve the quality of writing composition by helping students 'monitor their approaches to writing tasks.' They also argue that 'this type of instruction helps students focus on their goals, obstacles to reaching these goals, and strategies for overcoming the obstacles'.

3.3 Objective data, where available and purposeful, should be used; no apologies for using systematically and impartially gathered subjective data should be made. Branson et al. (1986:68) argue that evaluating educational programs is not easy because the collection and interpretation of evaluation data are difficult tasks. Furthermore, according to Williams and Bank (1981:131), 'the existence of test or evaluation data which reveal students' achievement in a particular subject or skill does not, inevitably, lead to beneficial changes in classroom instruction and concomitant increases in student learning.' In ESL programs, test findings are not often used sufficiently as data for evaluation, and the tests themselves are not always based on accurate or detailed data of students and programs. For example, Safi (1988:427) finds that 'the Secondary School Certificate Examination is unsuitable for admission to Kuwait University as the only criterion because (1) it is constructed annually from scratch and administered without pilot testing. There is no evidence of item analysis for reliability, validity and usability; and as a criterion-reference test it fails to develop norms for secondary school

graduates whose performance could be compared with a typical performance of other graduates in Kuwait and probably in the Arabian Gulf region.'

In other words, the identification of the various variables, and the appropriate data collection procedures are crucial to evaluation. Among the variables related to programs are books and other materials, classroom organization, techniques and sequences of teaching, and learning strategies. Data collection procedures relate to persons involved in the program, its processes and activities as well as its facilities. To gather all data relevant to evaluation, it seems natural to depend on both objective and subjective data.

3.4 Evaluation should be action-oriented with plans for moving reports to action. As it was noted earlier, ESL program evaluation differs in nature from other types of evaluation since the process of language learning is not only the means of achieving the aims of a program but also the goal of that program. Therefore, evaluation should be concerned with what actually goes on in the classroom as well as with what constitutes the content of the various components of the program: development of materials, use of media, coordination among program personnel, and the constant intercommunication and implementation of ideas. In this respect, KU PAE Project (1988) states that evaluation should be a continuous activity, or sequential procedures taking place at the various academic and administrative levels. Such evaluation should aim at preventing the accumulation of mistakes and at constructive interaction between staff members and personnel in various university sectors.

3.5 A plan for evaluating evaluation itself should be included. Evaluating evaluation should be personal and public. The latter could, of course, be designated by the institution to provide views on the overall evaluation act. The former is immediate, detailed and continuous. The personal evaluators of evaluation are those involved in the program (cf. Holt 1981:170). It is the teacher, for example, who is immediately most responsible for the effectiveness of his work, and he therefore should have his own say not only in evaluating his performance (e.g. in an institutional self-study) but also in terms of how he is being evaluated.

Related to the above issue is the need for systematic evaluation research (cf. Asamoah 1988) which should constitute a main component of a plan for evaluating evaluation. In ESL programs, the major issue is classroom-centered research. The process of second language learning and teaching, for example 'teacher feedback on learner error, teacher questions, turn-taking...interlanguage talk...' (Long 1984:422), has already received much attention. It remains to be seen how effective such classroom research has been in improving upon ESL program results. Furthermore, there is a need to study the factors that determine the impact of evaluation research findings on policy formation. 'These factors span every phase of the evaluation study, from its conceptualization through the dissemination and communication of its findings' (Siegel and Tuckel 1985:308).

Note

We are grateful to all directors, heads of departments, supervisors, and teachers who took part in the investigation and to all colleagues who read and commented on this paper.

References

- Al-Busairi, M. (in preparation). Needs, attitudes and motivation in foreign language learning: A case study of Kuwait University students studying ESP. Lancaster University dissertation.
- Al-Sani, M., A. Tawfik, I. Ashalabi, Y. Abdulmutti, and R. Alagha. 1981. Educational program evaluation in the Arab world. (In Arabic) Kuwait: Gulf Arab States Educational Research Center.
- Asamoah, Yvonne. 1988. Program evaluation in Africa. *Evaluation and Program Planning* 11:169-77.
- Becher, A., M. Eraut, J. Barton, A. Canning, and J. Knight. 1980. Accountability in the middle years of schooling. University of Sussex. (Quoted in Holt 1981.)
- Beretta, Alan. 1986. Program-fair language teaching evaluation. *TESOL Quarterly* 20.3:431-43.
- Branson, John D., M. Susan Burns, Victor R. Delclos, and Nancy J. Vye. 1986. Teaching thinking: Evaluating evaluations, and broadening the data base. *Educational Leadership* 22.2:68-70.
- Breen, Michael P., and Christopher N. Candlin. 1980. The essentials of a communicative curriculum in language teaching. *Applied Linguistics* 1.2:89-112.
- Cronbach, L. J., S.R. Ambron, S. M. Dornbusch, R.D. Hess, D.C. Phillips, D.E. Walker, and S.S. Weiner. 1980. Toward reform of program evaluation. San Francisco: Jossey-Bass. (Quoted in Nevo 1986.)
- Davis, Ed. 1980. Teachers as curriculum evaluators. Sydney: George Allen & Unwin.
- Dines, Peter. 1984. Examinations and evaluation. In: *Evaluating the curriculum in the eighties*, ed. Malcolm Skilbeck. London: Hodder and Stoughton.
- Dressel, Paul L. 1978. *Handbook of Academic Evaluation*. San Francisco, Washington, and London: Jossey-Bass.
- Gaffney, Jane, and Victor Mason. 1983. Rationalizing placement and promotion decisions in a major ELT program. *TESOL Quarterly* 17.1:97-107.
- Harlen, Wynne. 1974. Some practical points in favour of curriculum evaluation. *Journal of Curriculum Studies* 6.6:128-34.
- Holt, Maurice. 1981. *Evaluating the evaluators*. London: Hodder and Stoughton.
- Language Center. 1988. Users' report. (Unpublished). Kuwait University.
- Long, Michael H. 1984. Process and product in ESL program evaluation. *TESOL Quarterly* 18.3:409-25.
- Miller, Richard. 1988. Workshop on academic evaluation. (Unpublished). Kuwait University.
- Nevo, David. 1986. The conceptualization of educational evaluation: An analytical review of the literature. In: *New directions in education*, ed. Ernest R. House. London and Philadelphia: Falmer Press.
- Office of Assistant Rector for Scientific Affairs. 1988. Program academic evaluation project. (Unpublished). Kuwait University.
- Perkins, K., and Paul J. Angelis. 1985. Basic considerations for ESL program evaluation. *RELC* 16.3:72-92.
- Popham, W. James. 1975. *Educational evaluation*. Englewood Cliffs, N.J.: Prentice-Hall.
- Safi, A. Qayum. 1986. Kuwait University and its evaluation program. *Higher Education* 15:421-47.
- Saltzer, Milton G. 1982. The evaluation of an intensive English program. In: *The administration of intensive English language programs*, ed. Ralph Pat Barrett. Washington, D.C.: National Association for Foreign Student Affairs.
- Siegel, Karolynn, and Peter Tuckel. 1985. The utilization of evaluation research. *Evaluation Review* 9.3:307-27.
- Stufflebeam, D.L., W.J. Foley, W.J. Gephart, E.G. Guba, R.L. Hammon, H.O. Merriman and M.M. Provas. 1971. *Educational evaluation and decision-making*. Itasca, Ill.: Peacock. (Quoted in Nevo 1986.)

- Tyler, R.W. 1950. Basic Principles of curriculum and instruction. Chicago: University of Chicago Press. (Quoted in Nevo 1986.)
- White, Ronald V. 1988. The ELT curriculum: Design, innovation and management. New York: Basil Blackwell.
- Williams, Richard C., and Adrienne Bank. 1981. Uses of data to improve instruction in local school districts: Problems and possibilities. In: Improving educational evaluation methods: Impact on policy, ed. Carol B. Aslanian. Beverly Hills and London: SAGE Publications.
- Worthen, B.R. 1985. Program evaluation. International Encyclopedia of Education. 7:4090-96.
- Yalden, Janice. 1983. The communicative syllabus: Evaluation, design, and implementation. Oxford and New York: Pergamon Press.

Discourse frames and the cycle of instruction

Frederick J. Bosco and Anna De Meo

Georgetown University

It is convenient to think of conversations as cooperative performances in which, in Pirandellian fashion, the protagonists appear to write their own scripts. Like the characters in Pirandello's plays, we move about our lives creating our own scripts--some of which we fashion in negotiation with others; some we shape in our minds in anticipation of a situation or while reflecting on some past episode.

Imagine for a moment a situation in which you as an Italian parent are invited by your twenty-two-year-old son to a dinner at his apartment in order to meet his girl friend Elisabet. You've become accustomed to his enlarging circle of acquaintances and are not particularly moved by the prospect of meeting yet another one of his girl friends. Yet, things seem to be different this time. You sense that this relationship is more serious than previous ones. Your son's girl friend is Swedish, seven years older than he, divorced and with a two-year-old son. She's very attractive, sensitive and talented. She seems, however, a bit distant and taciturn. Perhaps your most serious concern is that she was previously married and has a son. And to your dismay, she is not Italian.

Many such situations in our lives call for responses of acceptance or rejection, confirmation or disconfirmation. Often we 'mask' our disconfirmation of a person or a situation in qualified terms. If you as the boy's mother, for example, were to express your feelings directly to your son, your approach might be very circumspect. If on the other hand you talk to your husband about the situation, you might use language that is much more direct.

We posed the above situation (hereafter called a 'discourse frame') to five dyads of native speakers (NSs) of Italian and asked them to simulate a mother-son and husband-wife interaction. Though reluctant at first, the NSs soon seemed comfortable in their new roles. We transcribed all their improvisations. Each transcription ran from one to two thousand words. We examined the organizational patterns of the spoken scripts, the speakers' choice of strategies, and matters of lexicon and grammar. In the process of analyzing the transcriptions, we began to ask ourselves a number of questions: If proficient performance is the goal of L2 instruction, shouldn't we use the same discourse frame with L2 learners that we used to generate NS performance? Wouldn't NS performances on videotape serve to provide cultural grounding and shouldn't the language 'chunkings' derived from such performances be used to channel instruction? We believed intuitively that the

answer to these questions was 'yes', and sought to work out a framework for examining and testing our intuitions.

Discourse frames and the instructional cycle. This paper presents the results of our investigation. We propose a strategy for organizing L2 instruction around 'discourse frames' in a three-phased process consisting of input, comprehension, and performance (see Figure A). The discourse frame serves within the instructional cycle as the fundamental, organizing 'unit' around which the various learning tasks cohere.

The term *discourse frame* refers to a situational context from which a number of interactive tasks (discussions, role plays, scenarios, etc.) can be derived. The frame itself consists of a mapping of a communicative state of affairs, including roles, shared background knowledge, possible agendas, and one or more interactive tasks (see Appendix 1).

The essential characteristics of the approach are:

(1) The basic input for the instructional cycle is derived from native-speaker performances. These performances are videotaped wherever possible for the benefit of L2 learners. One such performance (henceforth referred to as the 'scene') is selected as input in the instructional cycle.

(2) The NS performances are analyzed in terms of the agendas and verbal strategies employed, the information schemata and lexicon used, recurring syntactic structures and discourse markers. The language 'chunkings' resulting from the analysis provide the corpus of data to inform the comprehension phase of the instructional framework. During this phase of the process, L2 learners are taken through a series of microtasks to help them grasp the underlying routines, lexicon, syntactic patterns, and idiomatic expression which appear in the NS corpus of data.

(3) The same discourse frame which serves to elicit NS improvisation is also used in the instructional cycle to generate performances in the L2 classroom.

Thus, in our approach, the same discourse frame serves a threefold function: (1) to generate NS input, (2) to generate performance among L2 learners, and (3) to establish the basis for comprehension tasks in the instructional cycle (see Figure A).

The final section of this paper reports on the results of a pilot study in which L2 learners of Italian at Georgetown University were taken through the instructional cycle. Our initial study involved four dyads of L2 learners from our second-year Italian class. Two of the dyads of learners were taken through the instructional cycle in its entirety; the other two were taken through the discourse frame but did not have the benefit either of the model performance or of the transcription of the scene. All the preparatory work was done in Italian and outside of the regular class period. We wished to explore the question of what effect the instructional model would have on the performance of the students of Italian. We hypothesized that by way of our instructional framework, L2 learners would achieve a closer approximation to native-speaker performance in terms of appropriacy of register, cultural authenticity, and coherent discourse than is normally achieved through traditional classroom instruction.

Figure A. The instructional cycle.

**Phase 1
INPUT**

Introduce DISCOURSE FRAME. Ask students how they would handle the discourse frame. View videotape of NS interaction & introduce frame. Elicit possible agendas & strategies.

**Phase 2
COMPREHENSION**

Introduce representative language "chunks" from NS Corpus & promote tasks based on...

Verbal strategies used by NSs to achieve their goals. Information schemata derived from the discourse frame. Relevant lexicon, syntactic patterns, & idiomatic expressions.

**Phase 3
PERFORMANCE**

Re-introduce discourse frame. Modify one or more conditions. Rehearse for performance. Have selective students act out the interactive task. Discuss performances.

Set up new interactive task based on discourse frame. Act out the interactive task. Discuss performances.

Several lines of research converge in the instructional framework proposed in this paper: the important input studies of Krashen (1981); the comprehension focus of Winitz and Reeds (1975); the strategic interaction model of Di Pietro (1987); and the dimensions of comprehension and production in interactive language teaching of Wilga Rivers (1986). Our own study addresses the question of how to create the conditions for language use in the foreign-language classroom.

Phases in the instructional cycle. Consider the instructional framework outlined in Figure A. L2 learners are initially introduced to the discourse frame and given the opportunity to discuss various options inherent in the situation. The native speaker scene on video is then introduced and students have the opportunity to compare their reactions with those of the NSs. During the second or comprehension phase, L2 students are guided through the various complexities of the model scene. Various microtasks, focused on aspects of language usage and language use, are introduced to promote comprehension. Each task, whether message-oriented or medium-oriented, is directly related to the discourse frame in question. The third or performance phase recapitulates the initial input phase. The discourse frame is reintroduced and students are given the opportunity to rehearse for their own performance. In reintroducing the situation, one or more conditions should be modified in order to encourage innovation on the part of the students. In subsequent performances different agendas should be introduced in order to create what Di Pietro (1987:3) calls 'dramatic tension'.

NS improvisations as a source of language data. In order to gather a corpus of data for our instructional cycle, we worked with five dyads of NSs of Italian. Each dyad was presented with the discourse frame outlined earlier and was asked to simulate a mother-son or husband-wife encounter based on the communicative conditions outlined in the frame. Three pairs of NSs were videotaped during their improvisations. The other pairs were recorded on cassette tape. The resulting exchanges were transcribed and analyzed in terms of the types of agendas and verbal strategies employed, recurring lexical and syntactic patterns, idiomatic expressions, and discourse markers. There were both expected and unexpected results. One unexpected result was that in a number of the transcripts, there was not a single occurrence of the comparative construction. The subjunctive was used frequently. There were a number of phrases like : *io penso che tu sia ...*, *credo che*, *mi sembra che*, *non mi piace che*, *non sono convinto che*, all of which require an embedded sentence with the verb in the subjunctive. Conditional forms of the verb were also frequent: *vorrei che*, *mi piacerebbe che*, and so on. There were in addition many imperative constructions: *aspettate ancora un po'*, *cercate di conoscervi di più* and the like.

As expected, to promote the acceptance of the girl friend, NSs used various confirming strategies. Not only did they point out Elisabet's good qualities, but they alluded to those of her family. Consider the following sentences which illustrate a pattern of confirmation:

Allora mamma, che ne pensi di Elisabetta? È carina, vero?

È ancora giovane di spirito.

The topic of marriage is raised by Luigi almost from the start of the conversation. He announces to his mother:

... io so che le voglio bene, lei vuole bene a me e vogliamo sposarci.

Luigi's mother, however, tries her best to disconfirm her son's girl friend. She advises Luigi to postpone any thoughts of marriage until he has had more time to consider the consequences. Here are her words:

Sarà giovane finchè tu vuoi però ... però è una signora e è stata sposata per un certo periodo, ha un bambino, non è una ragazza.

... il fatto che lei viene da un ambiente diverso, da un paese diverso...

Però io continuo a insistere ... che sarebbe meglio ... che ci pensaste un pochettino ... e non far le cose di precipizio.

Some colorful idiomatic expressions occur in the NS improvisations including: *Ha un divorzio alle spalle, ha un figlio sulle spalle., ... sono rimasta da cani..., tu hai proprio delle idee così, un po' campate in aria.*

Comprehension phase of the cycle. In order to prepare the test groups of L2 learners for their performances, we took them through a series of tasks based on the linguistic and pragmatic content of the NS corpus of data. Two dyads of L2 students of Italian were presented with the profiles of Luigi, Luigi's mother, and Elisabet and were asked to describe each person from the perspective of one of the other participants in the situation. To illustrate:

Tu sei la madre. Cosa potresti dire della ragazza e di tuo figlio usando questi profili?

	Figlio	Ragazza	Madre
Nome	Luigi	Elisabet	Camilla
Nazionalità	italiano	svedese	italiana
Età	22	29	47
Professione	studente	pianista	insegnante
Carattere	allegro simpatico disponibile amichevole appassionato	sensibile affettuosa introversa riflessiva creativa	affettuosa possessiva espansiva impulsiva tradizionale
Aspetto	sportivo	attraente	elegante
Stato civile	celibe	divorziata con un figlio di 2 anni	sposata con un figlio di 22 anni

Mi sembra come persona.

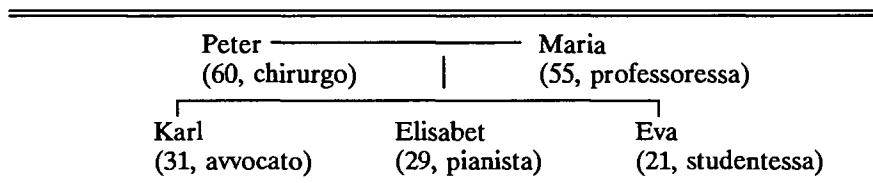
chiusa-aperta; riservata-sfrontata; introversa-estroversa; affettuosa-fredda

*Tu sei così ...
sicuro; disinvolto; allegro; disponibile; aperto*

In addition to the profiles, students worked with the following schema representing Elisabet's family (Figure B). The L2 students were asked to react to the following task:

Parliamo della famiglia di Elisabetta. Provate a dire qualcosa di positivo e/o qualcosa di negativo nell'ottica della madre.

Figure B.



Students generated various sentences which they also used in their performances.

Lei è colta. È un'insegnante di piano; ama la musica classica, è una vera artista.

Era nervosa di incontrarti. Sai che era la sua maniera di agire. È un po' introversa ma è molto affettuosa e sensibile.

Sai che viene da una famiglia colta. Suo padre è chirurgo e sua madre è una professoressa come te. Vedi come le due famiglie si somigliano.

Senti Luigi, io penso alla tua felicità. Loro abitano all'estero, non voglio essere lontana da te.

Characteristics of the NS and NNS performances. To illustrate the NS and NNS performances, we have chosen four brief excerpts as representative samples. The transcript for each segment is included in Appendix 2. Segment 1 is a sample of one of the native speaker performances. This performance served as the model in our instruction cycle. Segment 2 consists of an excerpt from one of the two control groups of L2 learners. Segments 3 and 4 consist of brief excerpts from our two test groups. All of the students used in the study were enrolled in the same section of Intensive Advanced Italian at Georgetown University.

Segment 1 (NS-NS dyad). The NSs in our study used frequent hesitations and repetition in their speech. Luigi's mother, for example, makes use of pauses and repetition as a strategy to express her displeasure with the entire situation.

NSs used a predominant organizational pattern that can be characterized as cyclical in nature. Every major point or concern is taken up two or three

times and at different segments of the discourse. Topics are not merely taken up once and dropped as they are in the NNS exchanges.

NSSs employed a variety of discourse markers and with great frequency. Such forms as *eh/ah*, *allora/dunque/beh*, *guarda/vedi*, *non so/forse*, *senti/ascolta* occurred frequently in the data. In the scene alone which consists of some 1300 words, the form *ma* (functioning as a discourse marker, not as a conjunction) was used 16 times while the form *eh* appeared 21 times.

Segment 2 (NNS-NNS dyad). This group of L2 learners did not have access to the NS scene on video.

This group, like the other control group, innovated much more than did the groups that had access to the NS performance on video. However, in so doing, they frequently broke out of the appropriate register. This did not happen with either test group. The control groups took advantage of various resources at their disposal, including literary texts which they had studied in class. One student, for example, in describing Elisabet used a sentence from Moravia: ... *è nel fiore dei suoi anni*. Though a poetic phrase, it does not follow the register required in the conversational exchange. Other such utterances included:

...*ha un'attitudine quasi innocente*.
...*e sa godere della vita*.

Elisabet is referred to as being *aristocratica* and *talentuosa*--words which are either inappropriate or have gone out of use. There were also some stereotypical expressions. In talking about Elisabet, one student asks:

Sa cucinare? Sa fare la pasta?

The not uncommon Hollywood stereotype of the Italian was evidenced on a number of occasions in the performances of the control groups.

Segment 3 and 4 (NNS-NNS dyads). These pairs of L2 students had the advantage of the video and transcription and were taken through the entire instructional cycle.

Both groups used several of the language 'chunkings' that occurred in the NS dyad. In each case, however, they made one or more lexical or syntactic change. The variations were in every case appropriate and grammatically correct. To give but a few examples:

NS *Che ne pensi di Elisabetta? È carina, vero?*
NNS *Cosa pensi di Elisabetta? È carina, eh?*

NS *È una signora, non è una ragazza!*
NNS *Lei ha 29 anni, è già una signora.*

Both groups used a number of idiomatic expression as did the NSSs in our study.

NS *Mi sembra che tu hai proprio delle idee così, un po' campate in aria.*
NNS *Guarda che le tue idee sono un po' campate nell'aria.*

NS *Quando finisci gli studi ti devi sistemare.*
NNS *Devi finire gli studi, devi sistemarti prima di sposarti.*

As indicated earlier, both tests groups remained within the appropriate register. This was evident by their choice of lexicon and the care with which they used idiomatic expressions. The cultural norms as established by the discourse frame were not violated by the test groups. They did not, for example, fall back on stereotypical behavior nor did they introduce extraneous elements in the situation. In sum, their interactions were culturally authentic in terms of the roles and the role expectations.

The organizational patterns of the NS samples differ markedly from those of the L2 subjects in our study. NSs returned to each topic on two or more occasions in the course of a single conversation; NNSs did not bring up any topic a second, third, or fourth time.

Conclusion. This paper has proposed a strategy for organizing L2 instruction around discourse frames utilizing NS input on video as a point of departure for instruction. Through a series of comprehension and interactive tasks, teachers have the opportunity to take the learner through the content of relevant schematic frames (agendas, verbal strategies, mappings of thematic content, role structure, and grammatical frames) in an orderly way. The NS improvisations provide a rich source of language chunkings and verbal strategies for the comprehensive phase of the instructional cycle. Teachers can also draw from the extensive pool of language data and analyses available in the fields of linguistics, sociolinguistics, and cultural studies. Schiffrin (1987), for example, furnishes extensive data on discourse markers. Hatch (1978), Long (1980) and others address the question of how discourse enters into the acquisition of a new language. Bosco (1985) shows that the content of utterances can be ordered by means of perceptual frames. Danesi (1985) reminds us that the results of contrastive analysis reveal much about how speakers of the target language use their culture to give shape to their conversations. Di Pietro and Bosco (1989) develop strategies to foster interaction in the classroom.

To conclude, we have argued that a firm focus on the discourse frame throughout the instructional cycle accounts for the integration of input, comprehension, and performance. Each of these phases involves particular activities: the input phase includes watching and assessing a video presentation; the comprehension phase incorporates tasks focused on language structure and the pragmatics of language use; and the performance phase involves acting out situations in the target language. It is the discourse frame that enables these varied activities to be shaped into a cohesive cycle of instruction.

Note

We would like to express our appreciation to all those colleagues and students who assisted us in our project. Special thanks are due to Robert Walsh, Roberto Severino, Bonnie Taylor, Camilla Tortorelli, Luigi Orlandi, David P. Harris, Lisa Yanguas, and Eugene Vricella.

Appendix 1: Discourse frame

Situation. Luigi is 22, the only son of Italian parents. His girl friend Elisabet is Swedish, in her late twenties, divorced with a two-year old son. Luigi arranged a dinner at his apartment to introduce Elisabet to his parents. Luigi's father couldn't make it to dinner because of problems at the office. Luigi's mother came by herself. Throughout the evening, she was not her usual, gregarious self.

Interactive tasks. The following day, Luigi has a chance to talk to his mother about Elisabet.

Role A. You are aware that your mother seemed somewhat cool toward your girl friend last evening. You are puzzled. Elisabet is very attractive and highly talented as a pianist. She comes from a good family. Her father is a surgeon; her mother is a university professor. You are thinking about marriage. It's important to you to have your parents' approval.

Role B. You've met your son's girl friend. She seems very much unlike your son. He is outgoing, intense, youthful; she is quite mature, less outgoing and not as affable. Besides, she's 29, divorced and has a two-year-old son. You've been very protective of your son and are uneasy about this relationship. How will you communicate your feelings about the relationship to your son?

Appendix 2: Transcripts from videotape (Excerpts from four performances).

Figlio = F.

Madre = M.

1. NS - NS

- F. Allora mamma, che ne pensi di Elisabetta? È carina, vero?
- M. Eh Luigi, son' contenta che ci hai invitati ...
- F. Anch'io!
- M. ... e quindi sono venuta a casa tua e questo m'ha fatto piacere, ma ...
- F. Ma?
- M. Ma ... ma sai, io ho l'impressione ... che lei sia molto importante per te adesso ... per cui ... forse ... tu non sei molto obiettivo.
- F. Ma, insomma ... io so che le voglio bene, lei vuole bene a me e ... vogliamo sposarci ...
- M. Volete sposarvi?
- F. E sì, sì!
- M. Ma sei matto?
- F. Siamo ... siamo convinti.
- M. Ma Luigi, ma ci devi pensare dieci volte. Ma tu non sai cos'è il matrimonio!
- F. Sì, ma ... abbiamo pensato molto e ... abbiamo pensato che è la cosa più giusta per noi due.
- M. Ma, più giusta. Ma, sai, il matrimonio non solo è una cosa seria, ma dovrebbe durare per sempre.
- F. Eh, durerà per sempre!
- M. Ma che basi avete?
- F. Eh, ci siamo ... ci conosciamo ormai da più di un anno e ... e andiamo molto d'accordo e pensiamo che possiamo anche vivere insieme sposandoci.
- M. E le differenze che ci sono tra voi due? Tu non le prendi in considerazione? Adesso non le vedi, sei innamorato, ma ... ma poi vengono fuori le differenze.
- F. Ma, le differenze possono anche aiutare un rapporto.
- M. E poi, innanzitutto, io sono rimasta ... sono rimasta da cani. Non me l'avevi neanche detto ... che è divorziata e ha un bambino!
- F. Sì, va be' ...

2. NNS - NNS (control group)

- F. Buongiorno mamma, stai meglio? Hai dormito bene?
M. Ma, abbastanza, ma ... ho perso un po' di sonno ... Sto pensando che questa ragazza ...
F. Ah, giustamente. Voglio sapere. Che pensi di lei? È meravigliosa, no?
M. Sì, è molto bella ...
F. È intelligente e anche talentuosa ...
M. Sì, intelligente, ma ... suo bambino [sic] è ...
F. Ah, è carissimo! Penso che mi vuole bene e anche io sto bene con lui. Non ha occasione di vedere suo padre e allora lo sostituo [sic] un po'.
M. Eh, giustamente, ma ... sei troppo giovane per questa donna e suo bambino [sic].
F. Ma, non è vero! Lei è ancora molto giovane. Ha un'attitudine quasi innocente, è nel fiore dei suoi anni ancora e ... e sa godere della vita. Allora non è un problema. Inoltre fra tu [sic] e papà ci sono 8 anni di differenza.
M. Ma, è diverso questo.
F. Perché? Perché nella mia situazione è la donna che è più vecchia? Ma mamma quello non è un argomento, non voglio ne sentire [sic] di più.

3. NNS - NNS (test group 1)

- F. Allora mamma, cosa pensi di Elisabetta? È carina, eh?
M. Eh, non lo so ancora. Ma tu sai, figlio mio, che mi è piaciuta così tanto la tua ultima ragazza, Francesca.
F. Eh, Francesca! Ma tu sai che Francesca e io abbiamo finito da due mesi fa [sic]. Perché parli ancora di lei?
M. Perché lei era perfetta per te! Era giovane, studentessa, di una famiglia molto bene e anche italiana. Ma questa Elisabetta è svedese!
F. E allora? Svedese, svedese ... queste sono cose che si dicono, non è la verità. Di più io ti dico che Francesca non era perfetta. Io lo so perché lei usciva con me e non con te!
M. Non lo so! Ma anche Elisabetta non è molto amichevole, eh? Non mi ha detto niente quando è stata qui!
F. Ma è perché non la conosci bene. Ma io sono sicuro che la conoscerai molto bene!
M. Perché? Cosa vuoi dirmi?
F. Vedi, mamma, noi abbiamo deciso di sposarci!
M. Cosa dici? Ma tu sei matto, eh? Lei ha 29 anni, è già una signora!
F. Forse, ma è giovane in spirito.

4. NNS - NNS (test group 2)

- M. E va bene, se l'amore è vero potete anche aspettare di più!
F. La conosco da un anno. Sono pronto per essere sposato.
M. Eh, ma sei pronto anche per essere un padre? Che farai con il bambino?
F. È un bel bambino! Io posso apprendere a fare da padre con il tempo.
M. Eh, non è facile essere un buon padre, come essere un [sic] studente. Devi finire gli studi, devi trovarsi un lavoro, devi sistemarti prima di sposarti.
F. Ma ha una famiglia ricca che può aiutarci. Sai che viene da una famiglia colta. Suo padre è un chirurgo e la madre è una professoressa come te. Vedi come le due famiglie si somigliano molto.
M. Si somigliano? Ma loro sono svedesi! Hanno delle idee differenti, sono liberali. Non sono come noi italiani.
F. Eh, ma forse tu sei un po' tradizionale.
M. Senti, Luigi, io penso alla tua felicità. Sei mio figlio unico [sic]. Loro abitano all'estero. Non voglio essere lontana da te!
F. Adesso non abbiamo l'intenzione di andare in Svezia.
M. Eh, va bene, ma questi svedesi ... Ho sentito che lei era molto fredda ieri sera.
F. Sì, ma era nervosa di incontrarti. Sai che era la sua maniera di agire. È un po' introversa ma molto affettuosa e sensibile.
M. Introversa? Ma tu sei troppo amichevole e simpatico per stare con una donna come lei!
F. È la mia giusta metà!

References

- Bosco, F. 1985. Pattern-practice revisited. In: *Toward an understanding of language: Charles C. Fries in perspective*, ed. P. Fries in collaboration with N. Fries. Amsterdam/Philadelphia: John Benjamins. 297-317.
- Danesi, M. 1985. Charles Fries and contrastive analysis. In: *Toward an understanding of language: Charles C. Fries in perspective*, ed. P. Fries in collaboration with N. Fries. Amsterdam/Philadelphia: John Benjamins. 277-95.
- Di Pietro, R. J. 1987. Strategic interaction: Learning languages through scenarios. Cambridge: Cambridge University Press.
- Di Pietro, R. J., and F. J. Bosco. 1989. The scenario: Its use in interactive L2 instruction. In: C. Faltis, ed., *Festschrift in honor of Robert Politzer*. In press.
- Hatch, E. M. 1978. Discourse analysis and second language acquisition. In: E. M. Hatch, ed., *Second language acquisition: A book of readings*. Rowley, Mass.: Newbury House. 401-135.
- Krashen, S. 1981. Second language acquisition and second language learning, Oxford: Pergamon.
- Long, M. H. 1980. Input, interaction, and second language acquisition. Ph.D. dissertation. UCLA.
- Rivers, W. M. 1986. Comprehension and production in interactive language teaching. *Modern Language Journal* 70.1:1-7.
- Schiffrin, D. 1987. Discourse markers. Cambridge: Cambridge University Press.
- Winitz, H., and J. Reeds. 1975. Comprehension and problem solving as strategies for language training. The Hague: Mouton.

Teaching language and culture: Old problems, new approaches

Ross Steele

University of Sydney, Australia

During the Round Table, I have observed that it is an established custom for American speakers to begin with some friendly chit-chat, with a joke or some humorous banter in order to put the audience at ease before beginning their paper. Coming from another national culture, I have to decide whether to imitate this behaviour pattern or not. In deciding not to imitate this custom, my intention is not to offend you. It is simply to attract your attention to the fact that many customs we take for granted in our own culture are not seen as natural by people outside the culture.

If my goal was to become totally assimilated into your culture, then I would try to integrate as many of your customs as possible into my repertoire of appropriate social behaviours. This integration process would be accompanied by various forms of anxiety. For example, I am going to deliver a paper at the prestigious Georgetown University Round Table; I have to begin with some humorous banter. How am I going to learn the appropriate tone for this banter? If nobody in the audience chuckles or laughs, then I have been unsuccessful in mastering the appropriate tone; I have failed in my aim to become assimilated, i.e. to be accepted as an insider in the group; I have been publicly rejected and so remain an outsider. This makes me generally more anxious and insecure in your culture. If my ability to get employment depended on my performance at this event, then I would be even more anxious and my sense of failure when you didn't accept me into your group by chuckling or laughing at my humour could have crippling psychological consequences...

This situation, with its general framework of desire for acceptance/success and frustration from rejection, is similar to the emotional situation we can unconsciously put our students in when, as is often the case, we slip into a teacher-centered approach (because of the institutional authority our position as the teacher gives us in the classroom) rather than working collaboratively and supportively with our students through a learner-centered approach to the foreign language and culture.

So, to avoid the anxiety of making a public fool of myself by failing to achieve the appropriate cultural banter before beginning my paper, I will, in keeping with the customs of my culture, say how honoured I am to have been invited to speak at GURT, to congratulate the organizers on the successful staging of the Round Table, and to thank you for coming to my session at the end of a busy day.

In other words, to get through a situation which is potentially threatening for my self-esteem, I have used an avoidance strategy. We frequently use avoidance strategies in everyday communication with each other.¹ However, in the teacher-centered foreign language classroom, students who don't give the expected answer are told they are wrong, lazy, ignorant, and so on. These students have perhaps used an avoidance strategy which reveals a deeper knowledge of communication strategies, a necessary step towards meaningful interaction. They feel frustrated when the answer which is valid for them is rejected without explanation by the teacher who insists on only one right answer. Their frustration increases and they decide not to re-enroll in the foreign language class. As Claire Kramsch says in her paper elsewhere in this volume, there are many voices in the language classroom and all of them have a role to play in the conversational narrative between the teacher and the students and between the students themselves. Even the silences have a voice. The narrative is already polyphonic when all the participants belong to the same culture. When the participants belong to different cultures, the teacher/conductor requires a huge amount of understanding, intuition, skill and crosscultural sensitivity to prevent a breakdown into cacophony and dissonant rivalries.

Until the audiolingual method of language teaching became popular in the 1960s, little attention was specifically given to the teaching of culture in a foreign language program. Until then the teaching of literature was the principal component of a foreign language program and the general aim of language instruction was to provide the learner with the necessary skills to read and appreciate, in the foreign language, the 'great works' of literature. It was generally considered that the grammar-translation method was the best way of providing the learner with these skills.

Culture was subsumed under literature. Literary masterpieces were seen as the finest expression of culture. At that time, the teaching of literature was done in an historical perspective. The life and times of the author were presented in detail, as was the dominant literary movement of the age such as classicism or romanticism. In fact these aspects were often given more importance than the actual analysis of the literary work itself. This analysis was usually a study of the themes rather than a textual study of the work as a unit of discourse. At the end of the course, students had thus received a general literary and historical culture which would give them access to a cultural elite and distinguish them socially above other citizens for whom the doors to the treasures of literature had remained closed.

The accelerating growth of the social sciences, particularly after the Second World War, gave increased public prominence to the disciplines of anthropology and sociology and the development of theories of culture. Literary culture was removed from its pedestal and described as only part of the total culture of a society. It was grouped with the other aesthetic expressions of culture such as the fine arts and music and called 'high' culture or 'big C' culture. However, these terms were not meant to imply superiority over 'small c' culture. In fact, 'small c' culture which gives meaning to everyday life has become, for anthropologists, the most significant expression of the behaviours, the attitudes, the beliefs and the values of a society. Access to 'small c' culture is through the language spoken by the society because

language has developed in response to the society's view of the world and its adaptation to the world. Language and culture are so tightly interwoven that they can never be separated--despite the conscientious efforts of generations of dedicated language instructors to focus principally on teaching grammar rules and exceptions to the rules! The word 'linguaculture' has been invented to stress the inseparability of language and culture.

The audiolingual method marks a turning point in the teaching of culture in foreign language programs. In his presentation of the audiolingual method, Nelson Brooks emphasized the importance of culture not for the study of literature but for language learning. He proposed 'a synthesis of culture as viewed by the scientist on the one hand and by the humanist on the other into an orderly and coherent program that can be meaningful in terms of the daily happenings in language classes at the earlier stages of instruction' (Brooks 1968:208). The social patterns for living, 'small c' culture, became the principal form of culture presented in the language classroom. Brooks did not exclude 'literature and the fine arts' and 'the sum total of a way of life' from the language classroom but considered it would be an impossible task to present the five aspects of his comprehensive definition of culture: (1) 'biological growth', (2) 'personal refinement', (3) 'literature and the fine arts', (4) 'patterns for living' and (5) 'the sum total of a way of life' (Brooks 1968:210). He therefore recommended that emphasis be placed on Definition Four ('patterns of living') 'with as much of Definitions Three and Five as can be reasonably added as the learner's competence increases' (Brooks 1968:212).

With the audiolingual method, language learning becomes a legitimate activity in its own right. Its main purpose is no longer to provide learners with the language knowledge necessary for their literature classes. Emphasis is placed on the spoken language rather than written language. The spoken language is introduced to the learner through daily life situations in the target culture. The learner thus becomes acquainted with behaviours and attitudes in the foreign society. However, the audiolingual method failed to fulfil the hopes of its proponents that learners would become fluent foreign language speakers. Stimulus-response techniques and constant repetition of structure drills did not enable learners to make the transition from controlled practice to free conversation.

The audiolingual method was gradually replaced by the communicative approach which aims to give learners communicative competence in the foreign language. Communicative competence was defined by John Gumperz in a paper delivered at this Round Table in 1981 as 'the knowledge of linguistic and related communicative conventions that speakers must have to initiate and sustain conversational involvement' (Gumperz 1982:325). In the communicative approach and its extension, the interactive approach which Wilga Rivers² has proposed as a strategy to focus on the exchange of meaningful units of communication between the learners and their interlocutors, 'small c' culture continues to have an important place. In textbooks using these approaches, it is introduced not only in the dialogues that usually begin each lesson and in the Cultural Notes that situate the topic of the dialogue in the target culture but also by the frequent use of authentic documents. In her analysis of textbooks, Claire Kramsch (1988) has brilliantly

demonstrated how cultural assumptions that are also implicit in the choice of topics and exercises are expressed on many discourse levels.

The use of authentic documents, which I define as texts produced in the target culture for a target culture audience, is justified on several grounds. They expose the learner to the 'real' target language used naturally for communicative purposes. Language is used to express meanings, not, as is usually the case in invented/pedagogical texts, to illustrate grammatical forms. Authentic documents contain a lot of redundancy which makes them easier to understand than invented texts in which this redundancy has been edited out. A variety of authentic documents provides an acquisition-rich environment in the classroom which can be compared to the natural acquisition environment in the foreign society. Because authentic documents present 'real' language, students find them more motivating than pedagogical texts. Authentic documents allow students to connect what they know from the world outside the classroom with the learning activity in the classroom.

Authentic documents can be selected from the print or the electronic media. At the early levels of foreign language instruction, authentic documents which combine iconic and written texts enable learners to transfer what they understand from the pictorial elements to decoding the meaning of the written elements. They are thus encouraged to apply to the foreign language the reading strategies they have learned in their native language. The teacher must, however, be on the watch for false cultural inferences which learners might make under the influence of their native culture and discuss with them why this is so. Simple advertisements, cartoons, forms that a tourist has to fill in, pages from telephone directories, radio and television schedules and short extracts from newspapers and magazines give beginners the opportunity to develop appropriate information-retrieval strategies in the foreign language. It is more difficult for them to understand recorded conversations or radio broadcasts where they have to rely solely on their listening skill. Contextualised prelistening activities which sensitize learners to the content of what they will hear reduce the anxiety they often experience when asked to comprehend a spoken authentic document. Combinations of audio and visual documents, such as video recordings and film strips or slides linked to a recording are other motivating ways to develop listening comprehension strategies. At the intermediate level, poems, short stories, short plays and television programs are authentic documents that can be used as language resources to provide learners with opportunities to increase the range of their communicative skills in the foreign language. Just as culture moved in the late 1960s from the exclusive domain of the literature class to the language class, so, in the 1980s, literature itself has been seen not only as an object for literary analysis but also as a rich source for a variety of language skill acquisition activities.

As other papers at this Round Table have pointed out, video is now part of our daily lives. Decreased production and purchase costs have made video authentic documents much more available for classroom use. Satellite dishes allow television programs from the foreign country to be transmitted directly into the language learning center and the classroom. However, the fascination of the screen must not beguile us into unquestioning acceptance of video as always superior to print. Video has the power to show us striking examples

of language use in an authentic cultural context. Comprehension of the spoken word is aided by visible paralinguistic cues. Different filmed locations can provide contextual elements for decoding linguistic exchanges and the sociocultural meaning of behaviours. However, is a televised news program necessarily superior to reading a newspaper? Does video by itself alter stereotyped views the learner may have of the foreign society? How much detailed attention do our students who are the products of the television generation give to what they are watching? What are their expectations? How has television changed the cognitive development and the learning styles of this generation? Have they all become skilled processors of visual information?

Video language courses have been produced in response to the perceived expectations of the television generation. However, despite their technical sophistication, they, like modern textbooks with glossy illustrations, can continue to transmit traditional attitudes and cultural clichés about the foreign society. It should not be assumed that video courses incorporate the advances that have taken place in language learning methodology. The currently popular 'French in Action' video course³ uses what is basically an audiolingual language learning methodology. Does the fact that it is a video course compensate for this and make it superior to a textbook course using an interactive methodology with video supplementary materials? The videodisc and the linking of the videodisc with computer technology provide further possibilities for learners to interact with the culture in formal language learning situations. Judith Frommer in her paper elsewhere in this volume gives a comprehensive range of language and culture learning activities made possible by these new technologies.

Authentic documents of many types from the print and electronic media have become readily available. This presents a new challenge to teachers who have usually not been trained as multimedia specialists. They have to be effective resource managers and select from the constant flow of authentic media documents those that are most appropriate for their students' learning needs in keeping with sound pedagogical priorities. This selection will depend firstly on the language being taught. If the language has a culture which is cognate with the learners' native language, such as Western European languages which are culturally cognate with English, learners will be able to call on their world knowledge and the interpretative strategies they have acquired in their mother tongue to help them decode authentic documents. I stress again the importance of the teacher's caretaker role during this decoding process to avoid learners transferring false assumptions from their own language and culture. If the language being taught has a culture which is essentially noncognate with English, such as Japanese and Chinese, different principles will guide the selection. The foreign script of print authentic documents will constitute an initial barrier between the English-speaking learner and the document. For these languages a convincing case can be made for first teaching the spoken language in association with video documents which show in context the behaviour patterns and social customs that are part of language use. Reading is not introduced until a later stage.⁴ It takes a considerable amount of time and effort to learn enough characters in the foreign script before the reader can comprehend relatively simple

authentic documents. Apart from recognizing words borrowed from English and icons associated with international products and services, learners of these languages, unlike learners of cognate culture languages, have few strategies they can transfer from their native language to help decode written texts.

Secondly, the choice of authentic documents will depend on the level, the needs and the goals of the learners. Communicative and interactive teaching is learner-centered. Instead of being a motivating factor for the learners, authentic documents which are consistently above their linguistic level become demotivating. The language of the document does not have to be exactly equivalent to the learner's linguistic level. If the document is slightly above the learner's level, the teacher can encourage the learner to adopt discovery techniques which activate what is already known in order for the learner to assimilate new language and cultural information and thus proceed to a higher level of competence in the foreign language. Recycling an authentic document at different stages during a course can be a way of motivating learners by showing them how their language skills have progressed. When the document is used at novice level, for example, learners would only be asked to perform simple tasks such as locating information (names, numbers, key words, etc.) and listing elements of information. The tasks to be performed with the same document would become increasingly complex for learners at more advanced levels, moving from partial to global to detailed processing of the text information. The diversity of language learner needs has been widely documented. Needs are influenced by cognitive and affective dispositions, individual learning styles, the learner's world knowledge and learner goals. Some students learn better with the help of print and computer materials, others with the help of audio and visual materials. Some students are learning the language for instrumental reasons, others for integrative reasons. Some students are only fulfilling a language requirement, others want to major in the language. Some students participate in study abroad programs. A range of these needs and goals will be represented by students in the same class.

Who are our undergraduate foreign language learners? In *International Studies and the Undergraduate* (a forthcoming publication by the American Council on Education), Richard Lambert publishes the results of a survey of the transcripts of approximately 9,000 students who had enrolled in foreign language instruction courses in 34 institutions across the United States (8 research universities, 9 comprehensive universities, 12 baccalaureate colleges and 5 community colleges). Table 1 shows the student attrition rate from elementary to intermediate to upper level language instruction courses for six languages: French, Spanish, German, Russian, Japanese and Chinese.

To take French, for example, in research universities, 46.3% of the students who enrol in French stop at the elementary level; only 41.9% of the students do the intermediate level and only 11.8% of the students take an upper level course.

Comparing these percentages with the other Western European languages at the research universities surveyed, only 10.2% of students take an upper level Spanish course, and 10.4% an upper level German course. The retention rate is higher for the less commonly taught languages: approximately one-fourth of the students take upper level Russian, one-third take upper level Japanese and one-fifth upper level Chinese. The initial number of students

enrolling in these languages is lower but the percentages show a greater tenacity by them to continue through the upper level sequence.

Table 1.

	University	Comprehensive	Baccalaureate	Two-year institutions
French				
elementary	46.3%	58.9%	40.0%	73.2%
intermediate	41.9%	22.2%	39.0%	23.2%
upper	11.8%	18.9%	21.0%	3.6%
Spanish				
elementary	45.3%	75.4%	52.6%	56.2%
intermediate	44.6%	18.9%	32.7%	42.3%
upper	10.2%	5.7%	14.7%	1.5%
German				
elementary	55.2%	56.9%	45.3%	76.5%
intermediate	34.4%	17.0%	31.4%	23.5%
upper	10.4%	26.1%	23.3%	...
Russian				
elementary	43.6%	27.5%	53.5%	100.0%
intermediate	30.8%	37.5%	26.7%	...
upper	25.6%	35.0%	19.8%	...
Japanese				
elementary	42.6%	80.0%	45.9%	100.0%
intermediate	23.4%	15.6%	29.4%	...
upper	34.0%	4.4%	24.7%	...
Chinese				
elementary	48.2%	52.9%	45.5%	60.0%
intermediate	29.6%	23.5%	20.5%	40.0%
upper	22.2%	23.5%	34.1%	...

SOURCE: Transcript analysis.

For the three other institutional categories a low percentage of students continue their language studies through the upper level although the percentage ratio from elementary to intermediate to upper varies between languages and the type of institution. In the baccalaureate institutions, there is a greater tendency for students to continue their language through upper level courses but even here the percentage of students is not large.

In summary, not only do these figures show the very low percentage of language students who continue their studies at the upper level, which has serious implications for this country's acknowledged international need for people with advanced level language skills; they also show that approximately half the students who enrol in a foreign language course at the undergraduate level only study the language for one year. In view of this, I believe we as teachers must ask ourselves what social and educational outcomes we wish to achieve and which linguistic and cultural outcomes we can realistically expect to achieve in a one-year course. It is illusory to think that in a nonintensive course, even with the most sophisticated technological aids, we can produce fluent speakers in one year. A chart drawn up by the Foreign Service Institute in 1973 which classifies French, German, Italian, Portuguese and Spanish as Group I, the easiest languages, shows that in a sixteen-week intensive course of 480 hours a speaker of English with minimum language learning aptitude will only reach a 1+ proficiency level, with average aptitude a 2 level and with superior aptitude a 2+ level. Arabic, Chinese, Japanese and Korean are classified as Group IV, the most difficult languages. In the same sixteen-week intensive course of 480 hours a speaker of English with minimum language learning aptitude will reach 0+ proficiency level, with average or with superior aptitude only a 1 level.

What then can be achieved in one year of undergraduate study of a language? Until now we have not been very successful in motivating our undergraduate students to continue to upper levels and the general public gives low importance to the acquisition of foreign language skills. Charles Ferguson and Thom Huebner (1989) report a similarly dismal situation in public schools. Citing Oxford and Rhodes (1988) that about five-eighths of secondary schools offer some foreign language instruction but in 71 percent of these less than half of the students are enrolled in foreign language courses, Ferguson and Huebner comment: 'This lack of commitment to foreign language instruction in public education is unique among industrialized nations (...) American educators give small place to foreign language instruction and on the whole do not expect students to acquire a working competence in the language they study; the brief exposure to a foreign language serves more as an inoculation against further study than as a foundation for achieving advanced levels.'

Given the high attrition rate of language students and the community attitude of low commitment to foreign language learning, I believe it is imperative to reassess our teaching goals and to demonstrate convincingly to the community the relevance of linguaculture by producing a larger percentage of the students enrolled in our courses who can use the foreign language with sociolinguistic competence. Reassessing our teaching goals would include reviewing the balance between language and culture in our courses. This balance should not be seen in terms of language or culture, skills or content, because just as skills cannot be taught without content, language expresses the world view, i.e. the culture, of the society that communicates through the language.

I propose that this review should have as its main point of reference an intercultural perspective. If we make our principal goal to give students crosscultural awareness and focus our language teaching within this

framework, the balance between language and culture will become a dynamic continuum which evolves according to the learners' needs. Learning verb tenses, for example, will not be an end in itself but, when appropriate, will be placed in the context of the society's concept of time. A comparison of the concept of time in the native culture and in the target culture will bring to the surface different cultural patterning in the two societies and reveal the different values their members attach to time in the rhythm of their daily lives and interpersonal exchanges. In the United States, punctuality is virtue and immediate results are expected in this product-driven society. In France, by comparison, you are not intended to arrive at the precise time given for a dinner invitation and intellectual discussions about long-term results are a valued part of the decision-making process. Yet the French National Railways are publicly proud that their trains, which transport a large section of the population, depart and arrive on time. Points of contradiction, such as this, within a society's system of values can be a stimulating starting point for exploring the network of cultural patterns that link the isolated pieces of information given in language textbooks. But let us not forget that such contradictions exist in the native culture as well as in the target culture!

The prevailing emphasis on 'small c' culture in foreign language textbooks has made learners better equipped to perform competently in travel situations. However, they often finish the elementary course with minimal 'big C' culture, the general cultural knowledge that the community and employers expect in a person having studied a foreign language.⁵ I suggest that our language graduates' lack of general knowledge about the geography, the main historical events, the social and political structures, and the major artistic and literary achievements is a significant reason why the community does not evaluate very highly what these graduates have gained from studying a foreign language. The community accepts that if you don't have the opportunity to use the foreign language you have learned, it is difficult to maintain your fluency in it. But it does expect you to have a general cultural knowledge about the foreign country, to be able to get information from sources in the foreign language when necessary, and to be able to give an enlightened opinion about cultural differences.

To remedy this situation, I do not recommend that some of the course time currently spent in language use activities should be replaced by lectures on 'big C' culture. The 'facts' given in these lectures are soon forgotten after the end-of-course examination. I would recommend, however, that some 'big C' topics be introduced into the language use activities and that learners be given the opportunity to discuss these topics from an intercultural perspective. I would also propose that students be trained in the skills necessary to locate and process information about the foreign country which they may need at some future time in their professional and personal lives.

In my view, our profession has placed too much emphasis on experiential learning in reaction to the grammar-translation-literature method of language teaching. A new balance needs to be found between experiential and cognitive learning so that our students become not only fluent travelers but also enjoy the intellectual stimulation that comes from exposure to other ways of interpreting society and the world. Teaching linguaculture in an intercultural perspective, we will be guiding our students not only through the different

stages of bilingualism but also through the parallel stages of biculturalism so that they become less ethnocentric and more accepting of cultural diversity.

Notes

I express my gratitude to the Mellon Foundation for granting me a research fellowship at the National Foreign Language Center at the Johns Hopkins University to continue my research on the teaching of culture in foreign language programs.

1. In Robinson 1988 chapter 7 there is a very perceptive discussion of the significance of avoidance strategies in the communication process.

2. I record here my personal debt and that of countless others in the language teaching profession, to Wilga Rivers for the many outstanding pedagogical insights in her writings and for her inspiring leadership role in our profession.

3. Pierre J. Capretz and Barry Lydgate. 1987. French in Action. Yale University Press.

4. Eleanor Jorden and Mari Noda use this approach very successfully in their interactive video course 'Japanese: The Spoken Language' (1987), Yale University Press.

5. A former officer with the Language Services Office of the State Department who had to recruit interpreters/escorts for official visitors invited by the United States Information Agency told me that many of the foreign language graduates applying for employment were unsuitable because although they had a specialised knowledge of literary vocabulary, they often did not have a practical day-to-day working vocabulary and, more importantly, they lacked essential background knowledge about the geography, the social structure, the political system and contemporary events in the visitor's country. They were thus unable to transpose what the foreign visitor was thinking and talking about into the American equivalents. An example was the incorrect equation of the U.S. Department of Interior with the French 'Ministère de l'Intérieur,' whose government function is responsibility for police and national security.

References

- Allen, W. 1985. Toward cultural proficiency. In: Proficiency, curriculum, articulation: The ties that bind. Northeast Conference Report. 137-66.
- Brooks, N. 1968. Teaching culture in the foreign language classroom. *Foreign Language Annals* 1:204-17.
- Ferguson, C., and T. Huebner. 1989. Foreign language instruction and second language acquisition research in the United States. Occasional Paper. National Foreign Language Center at The Johns Hopkins University.
- Gumperz, J. 1982. The linguistic bases of communicative competence. In: Georgetown University Round Table on Languages and Linguistics 1981. Washington, D.C.: Georgetown University Press. 323-34.
- Kramsch, C. 1988. The cultural discourse of foreign language textbooks. In: Toward a new integration of language and culture. Northeast Conference Report. 63-88.
- Rivers, W., ed. 1987. Interactive language teaching. Cambridge: Cambridge University Press.
- Robinson, G. 1988. Crosscultural understanding. New York: Prentice Hall. Second edition.

The role of language in the immigrant's life

Henry Kahane

University of Illinois at Urbana-Champaign

I was invited to describe the role of the 'other tongue' in the life of the immigrant who came to America. My native language is German, but I am not speaking to you as a teacher of German who knows the difficulties of the language and teaches his students how to overcome them. I speak, rather, about myself as an immigrant for whom the change from the first to the second language was, in all probability, the most incisive experience of his life. The linguistic adaptation has many aspects, and in the short time allotted to me I shall briefly point out to you the most significant ones, at least the ones most significant *to me*—in an ascending order: from the acoustic hurdles to the subcutaneous sociolinguistic implications.

The most elementary phase of the learning process is, as are all elementary experiences, the most pitiless: pronunciation. The pronunciation of the second language with the indelible habits of the first stigmatizes the poor man for his whole life as a foreigner. Two of the always present sins will suffice as examples. The interdental fricatives *th* and *dh*, which English does have but German does not have, are an eternal *memento mori*. Every single case of their occurrence requires a conscious effort which usually ends up with depression: the intended and beautiful *think* usually turns into a cheerless *sink*. And then the phonological rule of German which rejects the voicing of final consonants, contrasting with English, which admits them: *rib*, in the mouth of the German speaker, as long as he lives, turns into *rip*, and *kid* is *kit*, and *love* is *lof*.

If we climb up from phonology to grammar, the landscape (if it is possible at all) turns even more melancholy. I love to write English and even imagine that I write well, but certain weaknesses remain simply ineradicable, and even after half a century I still depend on the friendly help of the Salvation Army in weeding out those weaknesses of style.

First of all, the immigrant is entangled in a hopeless struggle with a giant called 'Tensesystem.' The present perfect is a nightmare: in German *ich bin gewesen* means that you are no longer there, but English *I have been* means that you *are* still there. This only slowly goes over into your system.

The worst for me is the position of the adverb, an almost hopeless problem, and the choice of the proper preposition to follow verbs (*think of* / *think about*). *Webster's Third*, always a source of spiritual comfort, unfortunately is too heavy to carry around in one's pocket.

As to the vocabulary, those of us immigrants who have come to their adopted country with what is called a 'humanistic education' are fortunate like spoiled children. The relatively limited Germanic constituent of English,

whose relationship with German is often not readily evident, vanishes in comparison with the colossal mass of Greco-Latin lexemes. These reached English partly with the Humanist Latinity, partly through the mediation of Old and Modern French, and they are easily identifiable.

But even this comfort, a well-deserved reward for school and education, is larded with obstacles. Words evidently common to German and English often develop quite different meanings and thereby turn into the so-called 'false friends,' who, so to speak, take us for a ride. Thus Germ. *eventuell* means 'possibly', Eng. *eventually* 'finally', Germ. *fatal* 'disastrous', Eng. *fatal* 'causing death', Germ. *Konkurrenz* 'competition', Eng. *concurrence* 'agreement', Germ. *kuriös* 'strange', Eng. *curious* 'inquisitive'.

In this context of a *semantic* distinction between identical lexemes, I should draw your attention to a dissimilarity of identical lexemes based on *stress*. The lexemes involved are the Latinisms and Latinized modernisms. The speakers of German stick rigidly to the Latin stress but when they speak English the familiar words, by exchanging the main and the secondary accent, become strangers. Germ. *Sùbstànz* turns into Eng. *sùbstànce*, and so it goes with innumerable technical neo-Latinisms: *Àspirin*, changes in America to *àspirin*, and a pharmaceutical youngster like *Týlenòl* would in German pronunciation instinctively come out as *Týlenòl*.

The entanglement of English and German forms has still another facet quite vexing to the immigrant: code mixing, in our case history, the insertion of English scraps into the German used among native speakers of German. Purists and teachers reject on principle this blending of languages. I must confess that I am an aficionado of it. I lived in many cultures and code mixing makes it possible to weave the idiomatic elements of 'the other' culture into whatever one's first language may be.

For the foreigners who smuggle English words into their language, these words reflect pragmatically the American milieu. It is the way in which, since time immemorial, borrowings have come into being. But for the speakers of German (and most of them are sinners in this respect) there is a catch in lexical code mixing. German nouns have a gender, which is reflected in the article, *der/die/das*, whereas the English article is condemned to genderlessness. But as soon as English words are inserted into German speech they need an article, which is dictated by the Sprachgefühl. After three days in Los Angeles one goes *zu der Beach*, one closes the window *auf der Porch*, one carries *das Luggage von der Attic*, for Thanksgiving *haben wir einen Turkey*. As a speaker of German I have not eaten, during the fifty years spent in America, a single *Truthahn*, but always *turkeys* with a masculine article.

So far, my remarks referred to the 'differences' between English and German. But I would like to draw your attention to a characteristic and interesting domain of the vocabulary which combines a German-English difference in the surface with an equivalence in the deep structure. I am speaking of the metaphor. Seen in a broad context, both languages share their metaphors with the other languages of Western Europe. Harald Weinrich (1976) defines Western civilization as a 'metaphoric community'. This metaphoric community is the result of a long historical process. The great contribution of the Greek mind consisted in its ability to abstract, i.e. to grasp spiritual events and to express them through the language of everyday

experience, i.e. paraphrase them metaphorically. Latin, behind which is always Greek (excepting the language of warfare) translated the abstract terms, and with medieval Latinity and the Humanist Latin of the Renaissance these Latinisms got into the Western languages. German plays its particular role in this quite significant West European event. In Germany, the repeated currents of a linguistic purism replaced the Greco-Latinisms by German 'translations', whereas English, which does not reject but accepts, kept the Latinistic vocabulary. Thus, the lexical surface (as I mentioned) is different in the two languages. Latinistic in English, nativized in German, but the deep structure is the same--the '*spiritus graecus*'.

A few typical examples from the hundreds which constitute our Western abstract terminology: Gr. *entýpo:sis* to Lat. *impressio* with Eng. *impression* but Ger. *Eindruck*; Gr. *perístasis* to Lat. *circumstantia* with Eng. *circumstance* but Ger. *Umstand*; Gr. *sympátheia* to Lat. *compassio* with Eng. *compassion* but Ger. *Mitleid*. Our grammatical terminology is Greek to a large extent: Gr. *ptôsis* is transferred literally into Latin as *casus*, which was kept in English as *case* but translated as *Fall* into German.

Whereas linguists are interested in a foreign language because it is a *language*, common mortals are eager to *use* the language, which opens to them a new civilization. Thereby, however, and without being aware of it, they get into the jungle which is summed up under the term *communicative competence*. This subfield, which is fast developing, aims at coordinating linguistic and social rules in a package deal. Let me illustrate this with two examples.

The German *du* (as the familiar counterpart of the formal *Sie*) has no exact correspondence in English. The first name of the person addressed is not a true English equivalent of *du*, which precisely with the 'change' from *du* to *Sie* reflects a change in human relations. In this country where your insurance agent after five minutes calls you 'Henry', the communicative value of the first name is low, indeed.

My second example refers to the enormous linguistic function of the 'number' in American life. Among the many committees of which, in the course of decades, I was a member, there was one which concentrated on the elite of whatever discipline was involved, and with every candidate the same question was put before the committee: 'In your opinion, is the candidate the best or the second best or the third best in his field?' I heard the words but I could not understand them. I am not able to arrange three good scholars (every one of them excellent in some specialty) according to 'first/second /third', as if they were running or jumping or swimming a race, where the 'first, second, third' can be ascertained by a clock--where, in other words, number is an empirical reality.

With two examples of communicative competence, the form of address and the belief in numbers, I hinted at the difficulties embodied in language, which the new immigrant into this culture must learn how to handle. My examples are certainly just two among many.

Let me end with a hint at the epistemological and ideological difference between the Old and the New World, which in my own professional life has been of considerable significance. What I am saying about myself has been

the experience of the many colleagues of mine who have gone through the same adventure of immigration.

My training as a linguist and, beyond that, as a so-called educated European, was historical. History and tradition provided all explanations. With the immigration into America I came from a diachronic into a synchronic world, which considers each stage as a system and explains it in terms of a system. It was an intellectual shock when I became aware of it. This was almost fifty years ago, but the shock continues to be felt. And in spite of serious efforts to think differently, my instinct still gropes for diachronic explanations of the world. But I learned to see also the other side, and for the historiographer of linguistics this is quite significant. Linguistics, like every humanistic science, reflects in the sequence of its phases the respective times and the prevailing cultures, and thereby becomes the key to those cultures--and this key is a blessing for the often perplexed and confused immigrant.

I described to you my personal experience as a European in America, and I meant 'European'. It is an old and often repeated observation that, seen against that extraordinary experience 'America', the ethnic differences, which are still alive between the various European cultures, fade away, and the Europeanism common to all turns into the prevalent characteristic. This observation is certainly right and very comprehensive--with one bittersweet and revealing exception, a shadow over which none of us can jump--our pronunciation of English.

New trends in foreign language teaching: Teaching English in the Italian medical school curriculum

Maria Ibba

Università Cattolica del Sacro Cuore, Roma

Abstract. The study of foreign languages, especially in the past few decades, has radically changed its basic objectives. It has shifted from learning the linguistic code (the language *per se*) in its formal aspects—phonological, morphological, syntactic ones—to the mastery of the language as a work tool and as a means of communication in specific fields of investigation and services. An ever-increasing professionalism, and a need for specialization due to rapid technological advances and to the increased and easier international scientific exchange, are mainly responsible for such a shift, together with some major socio-political-professional changes. Particularly sensitive to the emerging needs of the new type of learners have been scholars in the United States and in Europe. A phase of inquiry on the epistemological aspects of language for specific purposes was followed by phases of curriculum development and of ad hoc material production that have constituted valuable working aids, though not without the inevitable shortcomings of any emerging field of study. After several attempts in Italy, the Università Cattolica del Sacro Cuore has faced the issue of Language for Specific Purposes in a systematic fashion. It has organized congresses, symposia and workshops for the necessary theoretical discussions that were in some cases consequent, in others preliminary, to the addition of such new courses to the regular curricula. The teaching of English in the Medical School is one aspect of the above-mentioned new trend. Its goal is to meet the needs of those medical students and young doctors who wish to participate fully in the worldwide endeavor to further medical sciences. Some general as well as specific issues are dealt with briefly here, with particular reference to the work being done at the Catholic University Medical School, which is my institutional site for research and teaching duties.

1 Educational and political determinant factors for English language teaching in the medical school curricula

1.1 Educational factors. Medical research, like most scientific inquiry, is rewarded by continuous progress. It moves at times very rapidly in a succession of major breakthroughs and somewhat minor new findings, aiming at counteracting the threat of obstinate, serious diseases, at perfecting diagnostic and therapeutic procedures, or at slowing down the spread of some lethal epidemic.

It is well known that modern communications technologies are so advanced as to make any new discovery that takes place in any part of our planet potentially known to all other fellow scientists in the rest of the world in real time.

It would be desirable that progress in medical research, wherever it may occur, for the very fact that it regards the physical well-being of the human person, become as rapidly as possible the shared patrimony of everybody concerned in the medical world. Yet the latest technological advances in the field of communications are still a privilege for a few medical centers in the economically more fortunate countries to share their knowledge and expertise in cases, for instance, of exceptionally difficult diagnoses and treatment courses. The remaining, major bulk of communication takes place through regular channels, such as personal correspondence, periodical journals and congresses.

Written as well as oral communication of research results often takes place in English, regardless of the geographical setting of the study.

A good number of periodicals on the library shelves of Italian medical schools are written in English, even though many of them are published in countries whose official language is other than English. To mention only a few examples, four of the journals whose editors-in-chief are professors of the Medical School of the Università Cattolica del Sacro Cuore in Rome are written entirely in English, by mostly Italian authors.¹

The many international congresses, symposia and seminars that are held in the same University throughout the academic year always provide simultaneous translation, and one of the languages is inevitably English.

Highly sophisticated mechanical and electronic equipment for laboratory, diagnostic and therapeutic purposes always comes with instruction manuals for setup and use. The language of those manuals is usually that of the manufacturer's country, plus two or three additional languages, which rarely include Italian, but always include English.

An ever-growing trend is that of young medical doctors who set out to augment their advanced training with a study visit to some university hospital or research center in an English-speaking country.

Clearly, today's medical students and young doctors in Italy face a very complex professional challenge due to language throughout most of their training and practice, as well as in their research pursuits.

1.2 Political factors. In addition to the above-mentioned general language-related problems, the mastery of foreign languages has now become a compelling necessity for all citizens of the imminent, de facto European Community. As a matter of fact, the Recommendations on Modern Language Teaching² adopted by the European Council for Cultural Co-operation (CCC) at its 14th session (16-20 September, 1968) are based on the following principles:

that if full understanding was to be achieved among the countries of Europe the language barriers between them must be removed;

that linguistic diversity was part of the European cultural heritage and that it should, through the study of modern languages, provide a source of intellectual enrichment, rather than be an obstacle to unity;

that only if the study of modern European languages became general would full mutual understanding and cooperation be possible in Europe;

that a better knowledge of modern European languages would lead to the strengthening of links and the increase in international exchange on which economic and social progress in Europe increasingly depends; and

that knowledge of a foreign language should no longer be regarded as a luxury reserved for an elite, but as an instrument of information and culture which should be available to all.

Hence, in order for Italian doctors to be 'members in good standing', truly active members of the European Community and of the worldwide community of health professionals, the learning of foreign languages must necessarily be an integral part of their formal training.

Furthermore, the Italian government recently approved a new curriculum for all medical schools of the nation.³ This new curriculum includes the teaching of English as a compulsory subject. Also, the admission tests for entry in internship/residence programs include an English proficiency test, and English is taught during the first year of such programs and sometimes for a longer period.

1.3 An autonomous initiative. It has been six years now since the Università Cattolica del Sacro Cuore in Rome officially introduced the teaching of the English language in its medical school curriculum, thus starting a trend that other universities in Italy have been following to the extent that several full professorship positions in medical schools of state universities are already occupied.

Conferences, seminars and workshops on the languages of the sciences and on the new curriculum challenges they represent, have been taking place in Italy with increasing frequency in the past decade, and specialized publications are being produced in an attempt to meet a need that has finally been focused upon.⁴

The Università Cattolica's intense involvement in this new field of inquiry and its promoting and sponsoring many professional activities in such a pioneering field are made possible by two major factors: first, the ever keen interest of this institution in new issues, and second, its freedom to readily act upon newly emerged problems. This freedom, a privilege of the private university, is not enjoyed by state universities, which must await government authorization before any innovation becomes operative.

2 The experience at the Facoltà di Medicina e Chirurgia A. Gemelli in Rome. After an attempt to draw a general picture of the educational and political setting for the Italian curriculum changes and of some early implementations thereof, I shall now narrow the scope of this paper in order to answer briefly a few specific questions such as the following:

Who are the target learners?

What are their specific needs?

How is the teaching of the language being provided?

What is meant by 'teaching medical English'?

When in the medical school curriculum does the teaching of English take place?

2.1 Who are the target learners? The target learners are adults who partially fit the definition of 'adult student' given in the Council of Europe Modern Language Project document that reads: 'The term *adult student* refers to any person who has completed his compulsory schooling and who wishes to learn a modern language in addition to his normal occupation.'⁵

These adult learners are medical students and young doctors in their internship/residence programs. These two apparently similar groups actually constitute two different categories of learners with regard to teaching methodology, and especially to classroom and laboratory materials employed. Nevertheless, the final goal of the two groups is undeniably the same, inasmuch as they share the same professional interests and goals. A crucial difference between these two categories of learners is that while the students enter the language courses with no knowledge of medical matters, the doctors, instead, have all had their general medical training. Familiarity with medical matters, or lack thereof, determines to a significant extent the choice of teaching materials for the classroom as well as for the language laboratory.

A few words ought to be spent on pupils motivation and commitment. Like most adult learners who set out to acquire knowledge in some new field, our pupils are highly motivated to learn the English language, for the professional reasons stated above. As for most adult learners, the time to devote to the study of an apparently elective subject is hard to find; this is especially true for our learners who are so intensely involved in their medical studies and/or practice. Finally, like many adults who are ill informed with regard to what learning a foreign language is all about, our pupils often lack realism as to the amount of time, energy and persistence learning a foreign language as adults entails. A carefully prepared introductory lecture to the courses for students and physicians addresses this topic and provides the appropriate context for emphasizing the serious commitment the mastery of the language demands.

2.1.1 Student selection. Admission to the Catholic University Medical School is based upon the results of a battery of tests which assess aptitude, scientific skill and general education. Consequently, homogeneity of classes is guaranteed to a greater extent. Yet striking differences are found in the students' levels of proficiency in the English language, which is assessed in various ways, including the Michigan Placement Test.

For the English language courses the students are divided by level (only three at present, due to the limited number of specialized teachers). The intern/residents are also placed in the language courses according to the three levels above.

2.2 The learners' specific needs. Except for the difference between the two categories of learners mentioned in section 2.1, both groups share the need to acquire those skills in the English language which will enable them to engage successfully in various professional tasks that require the use of

English. Such skills range from the retrieval of some important information, through the reading of medical literature, to writing an original research report. In addition these learners will need to interact orally, with varying degrees of formality, with fellow doctors whose mother tongue is English or for whom English is the 'lingua franca' for communication with foreign colleagues. They will also need to participate in conference activities with lectures delivered in English.

All these skills entail at times a highly sophisticated mastery of the language. For instance, writing scientific discourse, either for a reader audience or for an oral presentation at a conference, requires familiarity with those devices that are employed to develop an argument, in a sequence of stages, to make up a stretch of language perceivable as a coherent whole.

Finally, as pointed out by learners with some experience, the relevance of 'social skills' in the English language needs to be taken into consideration.

An awareness of the students' needs was arrived at through three main routes: (1) my own reflection upon many objective needs that several years of work with medical students and doctors in Italy and abroad have enabled me to identify; (2) questionnaires administered to classes in the beginning stages of English language courses; and (3) questionnaires administered to experienced doctors not enrolled in the English language courses, yet very much aware of what a professional handicap a limited knowledge of English can be.

Of the three routes, the experienced doctors' input was the most enlightening one. The students and young doctors had only some general reference points, such as professional updating, study and research abroad, conference participation, together with the imminent reality of the European Community. The experienced doctors, on the other hand, were able to be very specific concerning the many language-related needs, including the sometimes underestimated, but undeniably crucial, 'social skills' in English.

2.3 How is teaching of the language being provided? In section 2.1 a distinction was made between two major categories of learners within the same medical school: students and physicians. Mention was also made of the possible methodological implications of such a difference and the consequent choice of teaching materials, as well as of classroom and language laboratory activities.

The students are all in the first or second year of medical studies.⁶ For the language classes they are grouped according to their level of proficiency in English--basic, intermediate, advanced--which is established by the administration of the Michigan Placement Test.

They all require 'General English', within the 'common core' range, in differing amounts. Remedial work is needed, together with the regular course work, due to the very different learning experiences the students have had. The considerably diversified groups call for much attention and flexibility on the part of the teacher.

Starting from a fair knowledge of the themes and topics of the medical profession, an inventory is being taken of the speech events in which the learners are more likely to engage. This inventory is by no means an exhaustive one, since a complete coverage of all possibilities is unobtainable,

but it seems to successfully trigger the learning process, whenever some use is made of those items.

Essential features of discourse, and especially the organization of the utterances of language as it is used in medical discourse, are some of the language clues included in the lesson plans for the physicians, together with some review of general language problems. The physician learners allow for the use in class of genuine, unaltered, nonmanipulated pieces of medical literature, at least from the intermediate course onward.

The materials experimentally used in such classes were collected in a 300-page manual which, temporarily at least, meets some of the students' needs. Those medical topics that proved to be of certain interest to the majority of the students were included in the manual.

The main criteria that guided the selection of the texts were the following:

- (1) Linguistic clues: Speech events the learner is more likely to have to use, and persistent morphosyntactic problems;
- (2) The specific context for the medical discourse: journal article, textbook chapter, lecture for oral delivery, etc.;
- (3) The medical theme: a topic that would interest as large a number of students as possible.

Another point of departure for classroom work is a continuously updated set of lectures which were previously heard 'live' at some international congress held in the home institution. They are replayed on videotape in order to focus on the organization of the discourse and its essential features of grammatical cohesion and semantic coherence.

Close observation of such models is followed by hands-on experience in the performance of some of the various tasks exemplified, to the point of rehearsing actual pieces of original work in class, with final destination the journal article or the oral presentation at some congress.

Within the time available I should like to briefly mention some of the main features of the manual. It begins with an introduction that aims at familiarizing the adult learners in a nonhumanistic faculty with the language sciences background for such a manual of English for the medical profession. As to the content of the ten units comprising the manual, it appears that, in general, the theme and the communication situation--such as the textbook chapter, the written research report, the lecture for oral delivery, etc.--determine to a considerable extent the prevalence in the discourse of certain linguistic features over others.

Here are some examples to illustrate this statement.

The unit whose text was taken from *Gray's Anatomy* provides a good context in which to observe operations, such as the definition and the description, to review expressions of spatial relations--the dynamic and static, the symmetric and the asymmetric ones--and finally, the use of the present simple of the verb can be revisited from various standpoints. All these are aspects of the English language that seem relatively easy for the opening unit but also extremely useful.

The unit dealing with future directions in diagnostic imaging lends itself particularly well to the discussion of how to make predictions, whether certain or possible, and how to express probability, proportions, etc.

Causal relations, derivational processes and the semantic field of drugs and medications is reviewed in the unit that deals with the treatment of cancer pain.

Implicit and explicit comparisons as well as the expression of possibility are found in the unit that discusses nodular thyroid diseases.

The unit which includes the text on the risk of myocardial infarction in cases of current or discontinued use of oral contraceptives provides an excellent model for the use of the passive form of the verb, which is predominant in the research article, a format usually quite synthetic and impersonal.

The list and/or elaboration of the sketchy statements above cannot continue lest time and space limits are severely violated.

3 I shall devote some time, nonetheless, to a brief discussion of some serious risks to which we would expose our learners by decontextualizing and depragmatizing beyond safe limits in an effort to describe too small a fragment of language.

The text that prompted consideration of this risk is the Declaration of Helsinki,⁷ which deals with duties and responsibilities of doctors involved in human research. Deontology is the theme of the text, linguistically the deontic modality prevails in it.

The attention of the class was focused mainly on the use the writer makes of the modals *must* and *should*. These two forms are usually taken to mean 'absolute necessity' and 'advisability', respectively. This is not the case in our text. Here, in fact, they are used interchangeably, only to avoid unpleasant repetition of the same verb.⁸ This is a case in which the instrumental function of modality prevails over the descriptive function. Therefore, unless the reader is aware of the writer's intentions and presuppositions, the moral and legal consequences of the misinterpretation in this case can be severe. Since the intentions and presuppositions of the writer/speaker are not always readily evident, it is the task of the teacher to guide the learners in the identification of possible given clues, in order to help them avoid the trap of ambiguity.

4 I will now finally get to the remaining two questions: What is meant by medical English? When in the medical school curriculum does the teaching of English take place?

Professionals in different fields such as physicians, lawyers, economists, etc., have always communicated among colleagues (about work) in a language that, to a good extent, is not shared by equally educated people of the same native language, but engaged in different professions.

These different 'languages of sciences' never constituted a special educational issue in any particular country, since studies and training gradually equip professionals with the specific language they need to use. But in the past two decades or so, rapid technological advances and the consequent disappearance of barriers between countries have caused the issue of the languages of the professions to emerge and impose itself on the professional

in language sciences. Thus, in the case being discussed here, defining the language of the medical profession became imperative for the language teaching specialist.

Before committing myself to any definition--that may not be reached in the desired, most precise terms--I should like to turn to the last question I set out to answer, the one concerning the point in the medical school curriculum at which the teaching of English takes place. We'll see presently why these two questions are better answered together.

Italian legislation set the time span during which the foreign language is to be taught in the medical schools as the first triennium. Returning to the question of what is meant by medical English, an answer might be the following: 'It is the English employed in the performance of various tasks by practicing physicians who keep up with the work of fellow doctors at home and abroad and share with them the results of their own professional experiences.'

This broad definition I have attempted of 'medical English' is intended to include all those communication situations mentioned in the learners' needs section of this paper: reading medical literature, interacting verbally with colleagues, writing research reports, presenting papers at conferences, and so on. All this entails mastery of what is usually referred to as the 'common core' of general English, as well as mastery of the many language devices and processes that concur to create pieces of coherent discourse.

Even without many justifying statements, I think that I can safely say that medical English as defined above cannot easily be taught in the early stages of the medical school curriculum, for it would be a thankless effort to try to get students to talk in English about matters with which they are unfamiliar. At best, medical practice-related vocabulary and basic science terminology can be introduced from the very inception of the course.

Thus, unless one ignores the complexities of medical discourse in the various communication situations and the importance of knowing the subject matter in order to be able to talk about it, the teaching of medical English ought to start gradually, with due respect for the learners' gradual acquisition of the 'res medica.' 'Words,' in fact, in their proper organization, come more easily if the object to be talked about is known.

What I have shared here reflects some of my experience in the past five years at the Catholic University Medical School. I have had to realize that progress is not always very rapid when venturing to work in a new field. Yet, I have opted for the optimistic attitude towards this challenge: rather than feeling lost and discouraged because of no sign of a path to follow, I have chosen the excitement of trying myself to make a path to walk on. It must be acknowledged that the many tools I have been using to make the new path are to be attributed to many fine craftsmen whose list would be too long for this context. I wish to mention at least one of them: the Dean of Languages and Linguistics of the Catholic University in Milan, Professor Sergio Cigada, who entrusted me with this challenging task and constantly provides valuable insights and encouraging support.

I am also very glad to take the opportunity to thank Georgetown University School of Languages and Linguistics, where most of my applied

linguistics studies took place, for the honor it has given me by inviting an alumna's contribution to its prestigious Bicentennial Round Table. Thank you.

Notes

1.. They are: Acta Medica Romana, Director Professor E. Manni, M.D., Dean; European Journal of Epidemiology, Director Professor A. Sanna, M.D.; Journal of Paediatric Neurosciences, Director Professor C. Di Rocco, M.D.; Rays, International Journal of Radiological Sciences, Director Professor P. Marano, M.D.

2.. J.L.M. Trim, R. Richterich, J.A. Van Ek and D.A. Wilkins. Systems Development in Adult Language Learning. A European unit/credit system for modern language learning by adults. Prepared for the Council of Europe. Pergamon Press Ltd., Headington Hill Hall, Oxford, OX3 0BW England.

3.. Decreto del Presidente della Repubblica 22 ottobre 1987. Gazzetta Ufficiale della Repubblica Italiana 15-6-1988. Serie generale n. 139. Articolo 23, Laurea in Medicina e Chirurgia, comma F): Corso di Lingua Inglese 'Lo studente dovrà seguire un corso di una lingua straniera, di regola la lingua inglese, fra quelle indicate nel manifesto degli studi. L'esame relativo, da svolgersi mediante colloquio e traduzione di testi scientifici, sarà effettuato entro il primo triennio.'

4.. The latest of these events: A Congress on the Languages of Specializations and Their Teaching in Secondary School and University, Brescia, 2-4 April 1987; The Proceedings: Il Linguaggio delle Scienze e il suo Insegnamento. Brescia: La Scuola 1988. In Fall 1988, workshops were held for teachers of the Lombardy region, on the Milan campus of the Catholic University.

5.. R. Richterich. Definition of Language Needs and Types of Adults. In: J.L.M. Trim et al. Systems Development in Adult Language Learning, Council of Europe Modern Language Project. Pergamon Press, 1980.

6.. The audience/readers are reminded of the fact that Italian medical school lasts six years, versus four years in the United States. Consequently, the early part of the curriculum deals with basic sciences, moving on to more specific medical subjects at some point in the second year.

7.. Declaration of Helsinki, Recommendations Guiding Doctors in Clinical Research. (1975) 29th World Medical Assembly.

8.. See the example: 'Biomedical research involving human subjects *must* conform to generally accepted scientific principles and *should* be based on adequately performed laboratory and animal experimentation.'

9.. These two questions are discussed at length in Ibba (1988).

References

- Bruner, J.S. 1975. From communication to language. *Cognition* 3: 255-87.
 Congress on the Languages of Specializations and Their Teaching in Secondary School and University. 1988. Brescia, 2-4 April 1987; The Proceedings: Il linguaggio delle scienze e il suo insegnamento. Brescia: La Scuola.
 Declaration of Helsinki. 1975. Recommendations Guiding Doctors in Clinical Research. 29th World Medical Assembly.

- Decreto del Presidente della Repubblica 22 ottobre 1987. Gazzetta Ufficiale della Repubblica Italiana 15-6-1988. Serie generale n. 139. Articolo 23, Laurea in Medicina e Chirurgia, comma F): Corso di Lingua Inglese 'Lo studente dovrà seguire un corso di una lingua straniera, di regola la lingua inglese, fra quelle indicate nel manifesto degli studi. L'esame relativo, da svolgersi mediante colloquio e traduzione di testi scientifici, sarà effettuato entro il primo triennio.'
- Hymes, D. 1971. *On communicative competence*. Philadelphia: University of Pennsylvania Press.
- Ibba, Maria. 1988. *L'inglese della medicina*. Vita Pensiero. Milano.
- Lyons, J. 1975. *Semantics*, 2 vols. London: Cambridge University Press.
- Richterich, R. 1980. Definition of language needs and types of adults. In: Trim et al. (1980).
- Trim, J.L.M., R. Richterich, J.A. Van Ek, and D.A. Wilkins. 1980. Systems development in adult language learning. A European unit/credit system for modern language learning by adults. Prepared for the Council of Europe. Oxford: Pergamon Press.

Lexical search strategies in L2: A developmental analysis

Irene Thompson

George Washington University

Introduction. This is an exploratory study of how beginning, intermediate, and advanced American learners of Russian cope with vocabulary limitations in a face-to-face oral interview. The topic of vocabulary search was chosen because it has not received much attention in L2 acquisition research although scholars have stressed the importance of lexicon in L2 acquisition (Hatch 1983:74). One particular area which has received some attention, however, is strategy use in lexical search. The study to be described is an attempt to pursue this line of research.

The study had two purposes. The first goal was to identify and describe communication strategies used by English-speaking learners of Russian when they were unable to retrieve a word or phrase to express their meaning. The study concentrated only on compensatory strategies, which Faerch and Kasper define as techniques '... aimed at solving problems in the planning phase due to insufficient linguistic resources' (1983:46). This definition is quite similar to the one suggested by Canale and Swain, who refer to communication strategies as devices that '... may be called into action to compensate for breakdowns in communication due to performance variables or to insufficient competence' (1980:30). No analysis was made of reduction strategies such as topic avoidance, message reduction, or message abandonment.

The second purpose of the study was to investigate whether there is a relationship between types of communication strategies used and learner proficiency level. It was expected that there would be observable changes in strategy use as a function of increased L2 proficiency and that examination of communication strategies of speakers at different levels of proficiency would yield data illustrative of the changing nature of pragmatic aspects of interlanguage.

Two previous studies provided a basis for the present investigation. A brief description of these studies follows. (1) Bialystok (1983) analyzed communication strategies of English-speaking learners of French who gave oral descriptions of pictures so that native speakers of French could accurately reconstruct them. The picture description task was an attempt to simulate a real communicative situation with a monolingual speaker of French. No help was given so that learners had to resort to strategy use when they did not know a French word or phrase. A taxonomy of strategies was developed according to whether the strategy was based on L1, L2, or was cooperative in nature. Bialystok found that less proficient learners used a significantly greater number of L1-based strategies than their more advanced peers. The

average number of strategies, however, was not related to proficiency level but the combination of strategies was. (2) Haastrup and Phillipson (1983) studied communication breakdowns of Danish learners talking to native English speakers and the resources these learners utilized in order to solve their problems. The researchers found considerable variation among the learners even though all of them made use of both L1-based and L2-based strategies.

The investigation described here intended to extend the findings of these two studies to Russian--a language somewhat further removed from English than French or Danish. The working hypotheses were: (1) learners will use the same strategies when facing lexical retrieval problems in Russian as they did as learners in the other studies, and (2) they will demonstrate increased reliance on L2-based strategies with improved proficiency in the target language.

Subjects. The subjects were 25 adult native English-speaking learners of Russian, five each from the following categories based on the ACTFL Proficiency scale: Novice High, Intermediate Low, Intermediate Mid, Intermediate High, and Advanced (ACTFL 1986). The subjects were enrolled in undergraduate and graduate Russian language courses at different institutions around the country.

Method. The sample consisted of tape-recorded oral interview tests conducted by experienced ACTFL certified testers. For purposes of this study the data obtained from oral interview tests had several advantages: (1) it provided a representative sample of interlanguage in a manner that simulated a real-life conversation, (2) it was elicited in a standardized way, (3) it contained a number of attempts by interviewees to engage in a variety of communicative tasks without help from the interviewer, (4) it represented a continuum of subjects ranked by actual speaking ability rather than years of language study.

The tape-recorded interviews were analyzed and relevant instances of lexical search were transcribed. Contextual information was noted down and all retrieval problems were identified on the basis of performance features that indicated that speakers were overtly experiencing retrieval difficulties. Performance features signaling difficulties in carrying out a communicative plan included different types of temporal variables such as slowed down rate of articulation, syllable lengthening, rising intonation, filled and unfilled pauses, containing either vocalizations such as *uh*, *er*, *hm*, or gambits such as turn-internal starters (*nu*, *vot* 'well'), cajolers (*znaete* 'you know', *znachit* 'I mean'), repetitions, and false starts. These time-gaining devices usually occurred not in isolation but in clusters.

It was assumed that not all lexical retrieval problems were marked by hesitations and other implicit signs of lexical search. No claim can be made, therefore, that the data collected contained all instances of lexical search attempted by the subjects. Performance variables offer only partial and indirect evidence about communication strategies because learners' use of successful strategies may pass unnoticed, particularly at higher levels of proficiency since advanced learners are better able to predict communicative problems and find solutions to them in advance. This makes performance

features less reliable as indicators of strategy use at higher proficiency levels (Rapauch 1983).

The communication strategies were analyzed according to a tripartite taxonomy, used by Bialystok and by Haastrup and Phillipson, which is based on the origin of the strategy as follows: 'L1-based strategies' referred to those instances when learners drew on their native English to solve 'lexical retrieval' problems, 'L2-based strategies' included those cases when learners relied on Russian to produce desired lexical items, 'cooperative strategies' represented speakers' attempts to involve their interlocutors in solving communication problems resulting from lexical deficiencies.

Results and discussion. Which strategies were used? The answer to this question was unequivocal: regardless of level, all learners used a variety of achievement strategies when they could not retrieve a lexical item. Numerous instances of the following achievement strategies were identified.

L1-based strategies

Language switch. Learners borrowed items from English without making an attempt to incorporate them into the phonological and morphological systems of Russian. For example: *Oni dali emu milliony dollarov, chtoby gotovit' ... uh ... guerrillas ... on president, on mog ... hm ... veto ... Congress mog veto.* 'They gave him millions of dollars to train ... uh ... guerrillas ... he is the president, he could ... hm ... veto ... Congress could veto'. The speaker simply inserted English words into a Russian sentence.

Calquing. Learners produced literal translations of English words or phrases. For example, *Oni konchayutsya ... kak skazat' ... uh ... vysokuyu shkolu.* 'They finished ... how do you say ... uh ... tall school.' Not knowing the Russian term for 'high school' the speaker translated it into Russian word-for-word.

Russification. Learners applied Russian phonological and/or morphological rules to English words or phrases to make them sound Russian. For instance, *On postavil Idaho na ... na ... hm ... na mapu.* 'He put Idaho on ... on ... hm ... on the map.' Instead of using the Russian word *karta* 'map', the speaker created a nonexistent word by adding the Russian accusative feminine ending */-u/* to the English *map*.

Language switch. As pointed out by Bialystok (1983:106), 'language switch' involves only L1, whereas 'calquing' and 'russification' entail elements of L2 as well, although the strategy itself is based in L1. Consequently these two strategies represent an intermediate step in the direction of L2.

L2-based strategies

Paraphrase. Learners produced descriptions of desired lexical items by listing their properties or functions, or by circumlocuting in a variety of ways. For instance, *Znaesh, yes' takaya meditsinskaya nitka ... prodayut ix zdes' v*

magazinax ... chistit' zuby ... v seredine zubov. ‘You know, there is this medical thread ... they sell them here in the stores ... to clean teeth ... in the middle of teeth.’ The speaker successfully described dental floss in terms of its function and where it can be obtained. A different set of devices was used in the following example: *Nu eto trudno potomu chto nado izuchat' vse eti ... eti... uh ... ya zabyla ... oh ... eti ... uh ... kanji nazyvaetsya po-yaponski ... no oni ne bukvy ... a ... kak risunki, no ne risunki.* ‘Well, it is difficult because one has to learn all these ... these ... uh ... I forgot ... oh ... these ... uh ... called kanji in Japanese ... but they are not letters ... but ... like ... uh ... like pictures but not pictures.’ The speaker had trouble retrieving the word *ieroglify* ‘characters’, and resorted to describing them in converse terms i.e., *ne bukvy* ‘not letters’, *ne risunki* ‘not pictures’.

Generalization. Learners produced an item which shared features with the desired item, such as a superordinate term or an item which was semantically related to the desired word or phrase. For instance, *Ona ... u nee yest' ... kariy ... net ne kary ... uh ... da ... karie volosy.* ‘She ... she has ...brown ... no, not brown ... uh ... yes... brown hair.’ The speaker overgeneralized a related adjective *kary* which can only refer to eyes in Russian. Or another example: *... i nado... goryachev solyonoy vodoy ... chisti ... chistit' gorlo kazhdiy den' pered snom.* ‘... and one must ... use hot water with salt ... to clean ... to clean ... the throat every day before bed.’ The speaker did not know the verb *poloskat'* ‘to gargle,’ and used the more general verb *chistit'* ‘to clean’ in its place.

Coinage. In place of desired words learners created nonexistent lexical items by selecting features of those items and incorporating them into the Russian phonological and morphological systems. For example, *U nego yest' dela s Yaponiey potomu chto on ... hm ... ne znayu kak skazat' ... on de ... on deiatel'.* ‘He has business with Japan because he ... hm ... I don't know how to say it ... he is a ... do ... he is a doer.’ The speaker made up a nonexistent word *deiatel'* ‘(literally:) doer’, by adding the nominalizing suffix *-el'* ‘-er’ to the verb *delat'* ‘to do’.

Self-repair. Speakers attempted to fix up trouble items by replacing them with others deemed to be more appropriate. For example, *Nu zhenshchiny mogut ochen mnogo dostat' ... do... do ... dobivat'sya v zhizni.* ‘Well, women can obtain ... achieve a lot in life.’ The speaker replaced the verb *dostat'* ‘to obtain’ with a more appropriate abstract verb *dobivat'sya* ‘to achieve’. This strategy was not always successful; for example, *Dlya menya eto bylo bol'shoe potryasenie ... ne potryasenie ... eto nepodtodyashcheye slovo ... bol'shoe ozhivlenie.* ‘For me this was a big shock ... not shock ... that is not the right word ... big animation.’ The speaker replaced the correct word *potryasenie* ‘shock’ with an inappropriate item *ozhivlenie* ‘animation’. This strategy, on the whole, was practiced mostly by more proficient speakers.

Cooperative strategies. All subjects used cooperative strategies, most of which took the form of different types of appeals to the interlocutor.

Requests for help. Learners asked the interviewer for a word; for example, *Ya soglasen s ego politikoy ... kak skazat' ... welfare program? ... programma pomogat' bednym*. 'I agree with his policy ... how do you say "welfare program?" ... a program to help the poor.' The learner first asked for help but then solved the problem by circumlocuting.

Requests for feedback. Speakers asked the interlocutor to evaluate the appropriateness of a lexical item they were not sure of. For instance, *Ya rabotal nad motorom ... ya ne znayu slova ... ya sovsem ego ... hm ... hm ... perestroi! Eto pravil'no? ... perestroi!* 'I worked on the motor... I don't know the word ... I completely ... hm ... hm ... rebuilt it. Is it correct? rebuilt?'

Admission of ignorance. Speakers indicated that they did not know how to say something in Russian, using this as a stalling device. With no help forthcoming, they often found a way to express the intended meaning. For example, *V vode mnogo ... pesticides ... ne znayu kak skazat' ... eto ... uh ... eto ximicheskie veshchi*. 'In the water there are many ... pesticides ... I don't know how to say it ... they are ... uh ... they are chemical things.' The speaker first switched to English, then admitted not knowing the Russian word, then finally resorted to a successful circumlocution. Another example was: *Ya plavayu ... ya zabyala kak po-russki 'breast stroke' ... uh ... uh ... brass*. 'I swim ... I forgot how to say 'breast stroke' in Russian ... uh ... uh ... brass.' The speaker successfully retrieved the correct word.

Is strategy choice dependent on speaker proficiency level? In order to find out whether the selection of strategies was influenced by the learners' speaking ability in L2, the number of occurrences of each strategy type was calculated for each proficiency group. The raw scores were converted to percentages in order to determine the relative frequency of each type of strategy. The results are presented in Table 1, which shows that as learners became more proficient in Russian they used fewer English-based and a greater number of Russian-based strategies to overcome their lexical limitations. This is consistent with Bialystok's study.

Table 1. Strategy use according to proficiency level.

Group proficiency level	Strategy type: L1-based	L2-based	Cooperative
Novice High	.63	.19	.18
Intermediate Low	.47	.31	.22
Intermediate Mid	.40	.40	.20
Intermediate High	.34	.43	.23
Advanced	.27	.49	.24

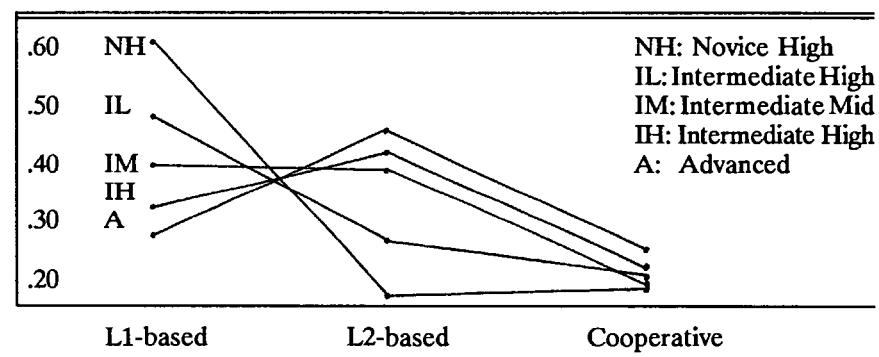
The predominance of English-based strategies among Novice speakers can be explained by the fact that they relied mostly on material that they had

studied. These speakers were as yet unable to use this material in an adaptive, creative way. As soon as they ran into communication difficulties they tended either to resort to reduction strategies or they switched to L1. The gradual shift in strategy use among Intermediate speakers from those based on L1 to those which operated with elements of L2 can be explained by the fact that these learners were attempting to use L2 'creatively', i.e. they were using what they had learned to deal with the communicative task at hand. By the time speakers reached the Advanced level, their strategic repertoire was characterized by predominance of L2-based devices.

This pragmatic shift points to a relationship between strategy use and degree of formal mastery of L2. As Bialystok (1983:117) pointed out, use of appropriate strategies '... is not entirely separate from the level of formal mastery achieved in that language, and the informal communicative use of language is constrained in this way by formal control over the system.' As far as cooperative strategies were concerned, their relative contribution to the overall strategic inventory remained quite stable across proficiency levels. This may indicate that socially based strategies are less influenced by formal knowledge of the language, particularly in interactive, conversational settings. It must be pointed out, however, that the complexity and variety of forms which were used to convey different kinds of appeals increased with improvement in proficiency.

The relative contribution of the three strategy types can be readily seen in Figure 1.

Figure 1. Strategy choice by learners at different proficiency levels.



While the results of this study were quite similar to those obtained by Bialystok (1983) and by Hastrup and Phillipson (1983), there was one noticeable difference. Bialystok's subjects used approximately the same average number of strategies across levels. However, in the present study, the average number of strategies was a function of speaking ability through the Intermediate-High level. This can be seen in Table 2.

This discrepancy may be related to method of elicitation. Bialystok controlled the amount of talk by having all subjects perform exactly the same picture description task. In the present study, however, the amount of talk

increased with improved skills in the language. On the average, Novice interviews lasted 11 minutes, Intermediate Low interviews 13 minutes, Intermediate Mid interviews 16 minutes, Intermediate High interviews 19 minutes, and Advanced interviews lasted 22 minutes. Thus, the longer the subjects talked, the greater the number of times they ran into problems and had to resort to strategy use.

Table 2. Average number of strategies at different proficiency levels.

Proficiency level	X strategies (N=5)
Novice High	14.6
Intermediate Low	16.6
Intermediate Mid	22.0
Intermediate High	29.6
Advanced	19.2

The drop in strategy use at the Advanced level can be explained by the fact that Advanced speakers with greater formal mastery of the language no longer needed to rely on compensatory strategies as frequently as their less proficient peers. Another possible explanation for decreased strategy use at the Advanced level may have to do with the fact that Advanced speakers could better predict possible communication problems and find solutions to them by planning ahead.

What is the blend of strategies within L1- and L2-based types? In order to answer this question, the relative frequency of each strategy was calculated within both types.

Table 3. Relative frequency of different types of L1- and L2-based strategies.

Strategy type	Proficiency level NH	IL	IM	IH	A
L1-based:					
Language switch	.72	.68	.52	.54	.18
Calquing	.06	.11	.27	.27	.54
Russification	.22	.21	.21	.19	.28
L2-based:					
Paraphrase	.31	.62	.36	.43	.47
Generalization	.15	.21	.27	.29	.24
Coinage	.54	.17	.30	.18	.11
Self-repair07	.10	.18

This breakdown shows that besides a shift from English-based to Russian-based strategies, overall proficiency in the language influenced the choice of particular strategies as well. For instance, within L1-based strategies, there was decreased reliance on language switch and increased use of calquing, although the use of russification remained essentially unchanged. Within L2-based group, paraphrase was the strategy of choice across all levels (except for Intermediate Low--a finding that is difficult to explain). Generalization tended to be used with about the same frequency at all levels. While coinage was the most frequently selected strategy for Novice speakers, its use declined at subsequent levels. Self-repair was not used by Novice and Intermediate Low speakers, and began appearing only at the Intermediate Mid level, with continued modest increase thereafter.

Summary. What do these preliminary observations on a small sample of learners suggest? In the somewhat stylized format of the oral interview, one can create conditions for eliciting speech and devise a taxonomy for classifying the variety of communication strategies that interlanguage speakers resort to when performance or competence limitations prevent them from carrying out their communicative plans. To what extent this could be generalized to real-life communication remains an unanswered question.

Spoken interactional data is extremely messy and difficult to measure quantitatively. It is often impossible to decide whether hesitation phenomena and gambits mark lexical search or whether they are indicative of some other cognitive operation.

In order to be certain that a speaker uses a certain communication strategy, we need to see multiple instances of that strategy throughout the sample. The empirical question is: how many instances are needed before it can be called a strategy? In the present study, some subjects did appear to favor a particular strategy. A good example was an Intermediate Low speaker who used language switch 11 times, or one Intermediate High speaker who relied on paraphrase to the exclusion of other strategies.

Despite the limitations of this small study, some generalization can be made. First, a simple taxonomy can adequately describe the lexical search strategies of learners at a variety of proficiency levels. Second, more advanced learners prefer L2-based strategies while less advanced speakers choose strategies which are rooted in their native language. Third, there are changes in the selection of strategies within both L1- and L2-based groups. These changes need further study. Fourth, the use of communication strategies is strongly related to mastery of formal aspects of language.

What are the implications for teaching and testing? As far as teaching is concerned, since it is difficult to separate strategy use from formal mastery of the language, it would be logical to conclude that strategies of communication should be addressed alongside other components of the target language. With regard to testing, if one were looking for watersheds that distinguish one level of proficiency from another, then perhaps the large decrease in L1-based strategies which occurred between Novice High and Intermediate Low speakers may be indicative of underlying changes in interlanguage. An analysis of such discontinuities may help shed further light on quantitative and qualitative changes at various stages of interlanguage.

References

- American Council on the Teaching of Foreign Languages. 1986. ACTFL proficiency guidelines. Hastings-on-Hudson, N.Y.: Author.
- Bialystok, E. 1983. Some factors in the selection and implementation of communication strategies. In: Claus Faerch and Gabriele Kasper, eds., *Strategies in interlanguage communication*. London: Longman. 110-18.
- Canale, M., and M. Swain. 1980. Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*. 1.1-47.
- Faerch, C., and G. Kasper. 1983. Plans and strategies in foreign language communication. In: Claus Faerch and Gabriele Kasper, eds., *Strategies in interlanguage communication*. London: Longman, 20-60.
- Haastrup, K., and R. Phillipson. 1983. Achievement strategies in learner/native speaker interaction. In: Claus Faerch and Gabriele Kasper, eds., *Strategies in interlanguage communication*. London: Longman. 141-58.
- Hatch, E. 1983. Psycholinguistics: A second language perspective. Rowley, Mass.: Newbury House.
- Raupach, M. 1983. Analysis and evaluation of communication strategies. In: Claus Faerch and Gabriele Kasper, eds., *Strategies in interlanguage communication*. London: Longman. 199-209.

Interaction and communication in the language class in an age of technology

Wilga M. Rivers
Harvard University

At times we need to pause, look back, and look forward to see where we are, to check the direction in which we are going and decide whether we are on the right track. In this paper, I do this, while incorporating ten principles of interactive language learning and teaching (Rivers 1989), which are listed in the Appendix.

The windup phonograph and the homemade crystal set with its BBC news broadcasts at 11:00 p.m. Australian time, heard distantly over what sounded like the waves of the sea--these were my first encounters with language and high technology. When we learned languages in school, at the age of eleven, there were no aural-oral aids, and most of our teachers' were nonnative speakers who had never been to the country where the language was spoken. (Ninety percent of us learned French for the obvious reason that Britain was across the Channel from France.)

Not that our language classes were silent--far from it. We sang (and danced as we sang), we acted, we played games, we practiced greetings, we drew and labeled vocabulary posters, and we recited poems. One day, in high school, a wonder appeared--we heard a Monsieur Stéphane (I'm sure that was his name) on a gramophone record. This was our first opportunity to hear native speech. He was attempting to convey French 'intonation'--that strange singsong! To us, he was singing the wrong song, and when our teacher tried to get us to imitate it, we burst into irrepressible giggles. At college (university to us), we had a native-speaking lecturer, a Parisian Armenian, M. Nazareth Karagheusian: a fabulous, vibrant personality who brought life and fun and overnight walks to the mountaintop to see the sunrise into our lives. Why would we need other aural-oral courses? Some of us went to Alliance Française *causeries* and chatted with retired Swiss jewelers and French railroad officials--the few who had made it to the Antipodes, and they were very few!

Soon, I was out in the countryside (the backblocks or Woop Woop) teaching in a town of three thousand, then later one of fifteen hundred (big centers in rural Australia). To keep up my French I talked to myself as I walked along the irrigation channels and kept a diary in French of my personal traumas. 'The cows don't speak French, Miss,' said the students. They asked me: 'Why do we learn French anyway? There isn't a France anymore,' since France had fallen--almost without a fight, it seemed--in the European conflict. However, my students continued to act out skits, sing and dance, and listen to quite long conversational-type stories in French. When

it could be made to coincide with the timetable, we tuned in to a radio lesson from the big city, where they could hear some authentic French speech. Meanwhile we avidly followed events in a Free French newspaper emblazoned with the Cross of Lorraine.

In the forties, news spread of a marvellous invention, talking tape--magnetic tape that could reproduce speech with great clarity and endurance would replace our old scratched, and scratchy, gramophone records. What a marvel! At last we would be able to have native speech in our midst at any time we wanted it, but realization of this dream was still a long way off. We had not as yet heard of the Army Method, pattern practice, or dialogue memorization (mim-mem). After all, plane trips were rare and no passenger ships sailed from Melbourne to far-distant parts from 1940 to 1949. In 1949 I traveled to England for France on the first ship to leave Melbourne after the hiatus, installing myself for the thirty-day voyage in a six-berth cabin over the engine--the cheapest berth on the Orion, paid for by patient saving over nine years. I was off to improve my knowledge of the language I was teaching in the country where it was spoken.

Where does all this lead? There was life before tape, cassettes, computers, laserdiscs, and CDs, and there was active, interactive, communicative language learning and teaching--not everywhere any more than there is now, but 'pockets of resistance', shall we say. We may have been short on native or even native-sounding models, but we were long on motivation, enthusiasm, and involvement. Interactive language teaching has a long history. (We may think of Montaigne learning to speak Latin on the tennis court.) So long as there is an imaginative, innovative teacher who encourages language learners to be active in comprehending and expressing messages in the language, language learning will continue to take place as it has since the first tribes came into contact with each other and bartered. No movement, approach, technological revolution, provision of or lack of funds can have a greater effect than the experience of learner and mentor in interaction, giving and receiving messages in an atmosphere of mutual respect and liking.

Then came the fifties. In Melbourne someone demonstrated a new approach using film strips and taped voices. CREDIF was taking its first timid steps into the Australian setting. After our free-wheeling approach to language learning, it seemed rather constricting--perhaps more perfect, but oh! so organized! so step-by-step, so meticulous! It was like taking a free spirit and clipping its wings, teaching it to step carefully and exactly--to dance like Skinner's pigeons.

Next was the language laboratory--a very important adjunct for foreign-language teaching around the world, though less obviously indispensable to those teaching second languages with native-speaking instructors in an environment permeated with the language. At long last, all language learners (including the instructor with little experience of native speech) could have access to authentic language, to the real thing. It was an undreamed-of luxury for many around the world, who had learned much by gluing their ears to staticky radio or had tried to follow foreign films, despite the frequent inappropriateness of the level of language.

What went wrong at that stage? The medium was there, but the inspiration was not. Those preparing the materials forgot that the students, the language learners, are not robots to be primed and set in motion, but individuals, each with his or her own interior motivation that determines what they will or will not endure. Somehow earnest scholars decreed that language learners would be stuffed with dull, tedious, and repetitious material, until they regurgitated it automatically, and not unexpectedly the students tuned out. Where possible, the students walked out of the language laboratory or refused to walk in. Materials were too frequently unrelated to what was going on in class or to what would be tested, despite the warning issued by Leon Dostert in the first Monograph of the *Georgetown Round Table* in 1951:

No matter how zealous the students, no matter how well installed the laboratory, no matter how excellent the equipment, the keystone to the effectiveness of the effort is what is put on the tape. The material on the tape should be completely integrated with the classroom work and should be indispensable to proper achievement in the course (De Francis 1951:77).

Only the most docile students would bother to conform to the situation that developed in many places, signup sheets notwithstanding.

Furthermore, the language laboratory became tied unfortunately to one methodology-one approach to language learning, with close ties to structural linguistics. Consequently, its fortunes tended to rise and fall with those of its originators, the practical applications of whose theories became identified, rightly or wrongly, with Skinnerian step-by-minimal-step programming. The forceful advocacy of structured aural-oral or audiolingual techniques had already raised hostilities in many schools and colleges. Teachers who had long achieved considerable success with modified grammar translation or direct method approaches, blended as often as not into an eclecticism that in their experience had worked well, now found themselves, at times roughly and brusquely, washed aside by the strong new wave that flowed into the schools. Those riding it forgot that teachers as well as their students are individuals, and just as varied in style and approach, with a need for confidence and the satisfaction of achievement in what they are doing. The simmering resentment of many soon expressed itself in forceful criticism of the language laboratory as producing mindless 'parrots', who could talk in the language while having nothing worthwhile to say, who could engage in language-like behavior, but not communicate meaningfully in actual person-to-person encounters.

There was and is nothing wrong with the language laboratory (now a multimedia center with cassette copiers, video facilities, computers, laserdiscs, recording rooms, film projection, and computerized lists of holdings). Nor, when it is well maintained, is there anything intrinsically ineffective about the hardware, which has been improving year by year, as have the layout of the center and its flexibility for offering all kinds of learning experiences. It was the courseware that was boring, poorly conceived, often prepared by persons other than teachers, who spoke the language but had little knowledge of language-learning theory and lacked pedagogical insight. Many of these materials preparers tried to put all students through the same series of hoops,

ignoring their individual needs in learning languages and their differing learning styles. Much of the material prepared, consequently, was useless as a supplement to what was going on in the classroom, where a revolution had already occurred. With a more active and communicative classroom the work in the laboratory seemed dull and irrelevant. Since academic rewards for putting effort into improving the situation were not forthcoming, many of the most interested instructors found themselves in a revolving door (a situation that could repeat itself in the computer age). Finally, as money became tight, the old materials were made to do one more and still one more turn on deteriorating equipment with fainter and fainter sound. Before there was time to remedy the situation (and in time there was a vast improvement in available materials) the word had gone forth that the language laboratory had 'failed'.

Of course, there was nothing wrong with the language laboratory. Used imaginatively and creatively, with regularly upgraded hardware and courseware, it provided unparalleled opportunities for individual learning, for vicarious contact with speakers of the language and their culture, for viewing as well as hearing, for producing as well as receiving. Students could engage in pseudo-communication (as near as technology could make it to the real thing), lying on their backs in the dormitory or jogging with their Walkman (Mueller et al. 1985). A generation to whom cassettes were a part of everyday life found nothing unusual in listening to foreign voices as they did to the music of other peoples. Materials became much more varied and demanding, as students listened to all sorts of authentic interviews, communicative interludes, serial episodes, soap operas, and talk shows that gave audio and visual access to life in other cultures. Such materials were available, of course, only when some instructor took the trouble to seek out or personally record them in the country or from satellite, prepare or acquire programmed tapes, and follow through in class with creative activities related to what students had worked on in lab sessions. As we know, every possible medium or modality can aid language learning, extending the contact far beyond the classroom walls. Nevertheless, expertise, imagination, creativity, and the will to relate what goes on in the lab or media center to what we know of language learning and student interest must be well in evidence. As Armand Bégué pointed out, at the second *Georgetown University Round Table on Languages and Linguistics* in 1951, 'the demands on the instructors are much greater with this type of teaching' (De Francis 1951:54).

We have come so far in language teaching, with the emphasis on communication, aspects of discourse, and analysis of language functions, as well as the creative aspects of reading and writing, that we must now be watchful. The entry on the scene of the computer and laserdisc leads to great expectations and great opportunities. If the expectations are too high to be realized we may miss out on the opportunities, drifting into a repeat cycle of boom-bust. It is easy to be naively dazzled by technology—with fantastic graphics and moving targets we can zap to make them disappear. Meanwhile inexpert programmers with few notions of what constitutes language learning design inappropriate courseware in an effort to compensate for the absence of language teaching specialists left behind in the forward surge. There is a real danger of a return to much drill and grammatical practice, with long explanations, which are relatively easy to program. These may be more

efficiently designed, providing more rapid immediate and individual correction than the teacher can give. They may provide the student with a more solid base of knowledge of the language but little opportunity to consolidate creative control for the expression of personal meanings.

If we slip back at this stage in language teaching it will take a long time to regain momentum. We already need to convince many reluctant colleagues that it is possible for students to learn to use the language in our classes, that we can provide a strong enough basis in experience for them to build on in the future. Many students, we know, return to language as adults to build on the foundations laid in one particular language or to profit from their experiences in language learning to take up a third or fourth language. In a 1988 survey of Harvard alumni in the classes of 1978, 1980, and 1983, 43 percent had studied at least one language following graduation and 14 percent had studied two or more, for professional, travel, or personal reasons (draft: Young Alumni Survey).

With these caveats, we may now consider the opportunities the present technological advance presents, while acknowledging the persistent problems in need of solution.

As we read of greater and greater novelties before us, with audio interactive video, CD-ROM, and a great range of international satellite emissions to be drawn on, we must keep in mind four basic considerations:

- (1) In what ways will the programs for use with the new technology fit in with the aims, content, and approaches of the courses they will accompany?
- (2) What can technology-based courseware accomplish as an aid to learning that cannot be achieved at lower cost (monetarily and in time and energy) in some other way?
- (3) How can we ensure that teachers wholeheartedly and advisedly cooperate in the incorporation of the latest technology into the language program?
- (4) Do we have research evidence that incorporation of the latest technological adjuncts leads to more efficient and effective language learning and language use?

Each of these questions needs to be answered with frankness and clarity of analysis before vast amounts of time and money are expended on hardware, software, and courseware. Questions such as these were continually raised in connection with the earlier language learning laboratory (Rivers 1981:425-28). They deeply interest decision-makers in the language-learning enterprise (administrators, fund distributors, and classroom teachers). Each will now be considered in detail.

- (1) In what ways will the programs for use with the new technology fit in with the aims, content, and approaches of the courses they will accompany?

To answer this question, we need to make clear to systems and program developers, as well as materials writers, the directions in which the profession has been going, these directions being influenced by societal pressures, not just teacher and researcher predilections.

There will be differences of opinion and emphasis at this point, but we can certainly say that ours is a period of diversity, where courses are developed to meet many objectives, practical, cultural, literary, or career-oriented. There is a strong emphasis on spoken communicative ability. Whether one begins with extensive listening or combines listening and speaking from the beginning, the objective is reception and production of meaningful and culturally appropriate discourse in normal contexts of language use. There is a growing stress on comprehending and interpreting written texts in culturally authentic ways, while producing texts oneself that reflect the discourse structure and semantic interrelationships of the new language. At all levels, students are expected to demonstrate a usable control of language for the purposes for which the course was designed. Pedagogically, there is a focus on learning through student directed and maintained tasks, cooperatively performed, that is, on the active language learner rather than the passive recipient of the wisdom of others. With this emphasis on control of language for whatever purposes, more attention is being paid to materials for advanced learners, a group whose needs were somewhat belatedly provided for in the language laboratory. Continual attention to materials for beginners may be easier but will ultimately be to the detriment of the profession who now have greater expectations of student achievement.

It is essential, then, that computer-assisted language learning set its sights much higher than the widespread drill-and-practice tutorials and grammar teaching, in which, in their common form, can be recognized the strong influence of uncreative programmed instruction. As Phillips puts it: 'We are busy pouring old methodological wine into the bottles of the new technology (in Leech and Candlin 1986:3). We need original thinking about new styles of programs that help students develop an understanding of how the language operates. Meanwhile, the most pedagogically sophisticated courseware producers have turned their attention to creative and exploratory interaction with the computer, which seizes students' attention and involves them with reception and production of language because of the intrinsic interest of the evolving situation, as in Brigham Young's Montevidisco, MIT's Athena project, the BBC's Domesday Project, and the Cambridge Micro Software Granville package (Smith 1989: 239-43; 99-107; 171; 166-67).

If CALL is to realize its professional potential, we must keep emphasizing the need for innovative thinking in the production of a diversity of materials that promote creative use of language by the learner. The latest technology is ideally suited to today's eclectic atmosphere because, of themselves, the computer and laserdisc are neutral; a well-informed programmer can adapt them to any requested style, content, or approach.

(2) What can technology-based courseware accomplish as an aid to learning that cannot be achieved at lower cost (monetarily and in time and energy) in some other way?

The major advantage of technology-assisted language learning is indubitably its potential to meet the needs of individual learners, giving students time and the opportunity to repeat or not repeat material at will, providing immediate access to the desired information, illustration, correction, modeling, or guidance, or the right to ignore such aids. The student may

choose more or less practice, exploration, advice, or evaluation. *Within the limits of the programming*, students may follow their own bent in language learning. The words 'within the limits of the programming' are crucial here, since this potential can be realized only if the courseware designer and programmer are alert to language-learning needs and provide these kinds of opportunities. As Hope, Taylor, and Pusack (1984:16) have observed:

Though virtually tireless, computers can become tiresome. If a student gives a creative answer to a grammar exercise, a classroom teacher can respond appropriately, laugh or make an acknowledging gesture, and come up with a good rejoinder. The computer program typically searches for patterns and, finding none or few in an imaginative or witty response, produces a standard error message. This kind of behavior might discourage creative efforts.

If we accept the principle that 'basic to use of language are knowledge of language and control of language' (Appendix:5), that each language user has a mental representation of how the language works, an individual competence that governs the construction of utterances, then we see a place for students working to increase their knowledge of their new language. This does not necessarily mean teaching and drilling of traditional rules of grammar. This knowledge of language comes through performing rules in meaningful material, with conscious focus on difficult areas. With thought and care the computer can be programmed for what is essentially individual effort on the part of the student (some assimilate this knowledge faster, some more slowly), thus liberating the classroom for more creative interactive activities that stimulate communication through language. Otto (1988:84) has observed that 'learners are willing to practice if such experiences build confidence and competence at the "performance" stage' (1988:84).

The question arises whether with interactive programming the computer can also provide for the student the necessary experience of communicating meaningful messages in the language, which is essential for developing control of the language. To date this has had to be limited to keyboard responses, except where the computer is trained to recognize certain oral responses and match them with a template of acceptable production in the language--presently a very limited procedure. Even with the more sophisticated speech recognition the future may bring, we have to admit that communicative episodes with a computer are still in the realm of pseudo-communication. The computer is programmed to pick up certain words and expressions and react to these as some therapists do, in the following way: 'I'm having trouble with my husband'; 'Tell me about your husband'; 'He never talks to me over breakfast'; 'Is breakfast a problem?' (along the lines of: find A, react to A; find B, react to B). The human participant is often amazed at the interactive potential of the computer in what is basically a prompted monologue.

The need for more creative interaction in CALL has been taken into account by some in the provision of imaginative, task-oriented games and simulations where student decisions affect the turn of events, graphics respond to student directions, and something like real-life interaction appears to take place, even if still mainly in the form of responses on a screen. Davies points

out that such programs take a long time to write (in Smith 1989:167); they are, therefore, not as numerous as teachers would like. Practical use has shown that these types of programs provide for genuine communicative interaction when students work together in groups at the workstation, the challenges of the CALL activity stimulating them to lively discussions, disputes, and cooperative decision making (Higgins and Johns 1984:64; Smith 1989:99, 166; Young 1988).

It may be objected that computer games and simulations give mainly comprehension practice, particularly of written texts, although an audio adjunct is becoming more generally available, and that simulations can be performed in class (e.g. the evolution of a border dispute between two imaginary countries). However, with the laserdisc, students can enter into another environment and see opportunities for action, while the interactive program varies the information available in accordance with the direction of student decision making. The answer to this objection is the combining of live interaction with the interactive program, either at the workstation or in the class group.

Whether CALL realizes its potential for interaction is dependent on recognition by materials developers of the essential ingredients of interactive language learning and their pedagogical skill in incorporating these ingredients into the materials, as well as on the ability of the language teacher to incorporate work with creative CALL materials into classroom activities in ways that maximize their contribution.

We come now to the teacher.

(3) How can we ensure that teachers wholeheartedly and advisedly cooperate in the incorporation of the latest technology into the language program?

Writers on CALL often cite the fear teachers have of being displaced by computers; some are still hiding behind the classroom door. This was also one reaction to the introduction of the language laboratory, yet that fear was far from being realized. More serious is the feeling many fine teachers develop that, after years of recognized success in the classroom, their work is now being underrated in favor of an impersonal mass of plastic and silicon. Willing as they may be to move with the times, they anticipate that to incorporate this new medium into their work will require much time and effort for which they may not be adequately recompensed, and that, without special provision, it will eat into already meager funds for supplementation of teaching. Classroom teachers need to be made aware of the user-friendly qualities of available authoring systems for preparing CALL materials like Frommer's MacLang and Pusack's Dasher, and the availability in language-teaching journals of evaluations of the relative usefulness or effectiveness of commercially available courseware. Then there is the well-known propensity of the computer to crash or 'go down' at inconvenient moments. (Didn't tapes break or get accidentally erased, and weren't films sometimes wound backwards? Teachers murmur, '*Plus ça change...*') As happened with the language laboratory, teachers suspect that institutions may install equipment, while being very parsimonious about supplying funds for a variety of materials

and for skilled technical help in looking after the enterprise; they must see that a budget is established that provides for more than the initial installation.

The problems teachers foresee are very real for them, since so few of those presently teaching languages have had any training in incorporating CALL into their courses. Teachers need to be fully involved in decision making when consideration is being given to the introduction of CALL equipment, so that from the early stages they have begun to think about how they will use it. Above all, they should be given hands-on training with available materials well ahead of implementation with classes--training which gives them the opportunity to experiment with authoring systems as well, so that when implementation occurs they will be working with materials they feel are appropriate for their students and that bear the mark of their individual style of teaching. The occasional workshop at a professional meeting is insufficient if present approaches to teaching are to be advanced by computer and laserdisc programs.

Without the cooperation of well-informed teachers, precipitous introduction of new technology can lead to equally precipitous abandonment of what may be prematurely judged to be an ineffective enterprise. Teachers must analyze which parts of the teaching endeavor are advanced by technology and which can best be left to their own personal interaction with students. Much 'inauthentic labour', to use Kemmis's term (Phillips in Leach and Candlin 1986:4), can be taken over by the computer (that is, revising and correcting of written scripts, drafting tests from banks of items and evaluating them, keeping individual totals of correct responses, and other chores that are by-products of the language-learning scene). Teachers can be readily persuaded to yield domain in this area. Many teachers, on the other hand, believe they can adapt more flexibly to student needs than a programmed machine does. They should consider carefully whether this is actually the case and study how their students can profit from the best of both worlds. Since ultimately only the teachers themselves will ensure the successful incorporation of CALL into students' language-learning programs, consideration of their viewpoint is vital.

(4) Do we have research evidence that incorporation of the latest technological adjuncts leads to more efficient and effective language learning and language use?

The field of CALL is still very young and longitudinal studies of learning take much time to design and complete. Experience has demonstrated that in large-scale experiments too many variables and unforeseen events come into play to produce more than 'final results of questionable validity and interpretability' (Clark 1988:19). Hope, Taylor, and Pusack speak of 'inconclusive empirical evidence concerning the effectiveness of CALL' (1984:9) at the present time. Clark considers that microstudies, like Robinson et al.'s 1985 CLCCS investigation, are more useful at the present stage of development than macroresearch (1988:17-20). Robinson's study found that:

In organizing material for CALL...lesson presentation, language-learning materials may be more effective, over time, when presented and practiced within an *integrated context* in which students' attention is focused on the

meaning of the material and language is used to *draw inferences* as in *solving a problem*. Material may be more *meaningful* when students *relate personally* to it, either because the materials contain *reference to themselves* or to *people they know*, because it is amusing or otherwise *emotionally appealing* and because they *select it (from a menu)* out of *personal interest*. While these features did not appear to have any *immediate* effect on second-language learning...their *cumulative* effect is noteworthy (Robinson in Smith 1989:131).

This study also investigated forms of error correction and found that various ways of leading students to discover the correct answer for themselves (for example, highlighting errors and blanking out wrong characters, while giving hints; rephrasing the question to give implicit correction through modeling, or following up an incorrect response with a parallel type of item) were more effective than merely displaying the correct response, with or without explanation (*ibid*:131-32).

Another microstudy of Young found that programs that 'allow the learner to negotiate the outcome of the activity' generated among students at a work-station discourse that was 'closer to ordinary conversation' than programs where the students merely decided what to fill in or rearrange on the computer screen (1988:65-83).

A USAFA study found that active involvement of the student through mediated interaction on the content and language of filmed material on videodisc, which included feedback on incorrect choices on content with hints, vocabulary lists, and video replays, led to dramatically better results in achievement and retention, as measured by posttests (Verano in Smith 1989:257-61). Other useful studies can be found in journals and books on CALL.

Experimental courses are being developed that bear watching. The Harvard-Stanford Collaborative Newspaper Writing project, funded by the Consortium on Language Teaching and Learning, enables intermediate-level classes, under J. Frommer and J. Barson, to produce cooperatively a newspaper in French, by desk-top publishing via modem from coast to coast. This is purposeful, task-oriented language learning, through which students acquire career-related skills of computer and word-processor use, as well as the vocabulary and expressions related to editing and desk-top production. The cooperative nature of the enterprise stimulates student motivation through the element of competition.

Teleteaching, similarly, is enriching distance learning, as in the Annenberg/CPB project in science and technology which links American and Chinese universities. Teleteaching has great potential for bringing real contact between learners of a foreign language and speakers of that language in their own land. Courses have been developed that link business students and foreign-language learners in projects where the former must depend on the input of the latter. The Athena project (Annenberg/CPB Foundation funded) at MIT takes students on investigative adventures in Colombia, which twist and turn according to student decisions, and to Paris, where once again student decisions appear to affect the lives of the protagonists of the program. Meanwhile, poltergeists upset a German home that students must reorder in

effective German (Murray, Morgenstern, and Furstenberg 1989). Innovative courses await only imaginative enterprise in exploiting available technology to be brought into being.

Video and computer can be enlisted in the study of language-learner behavior, as Jorden has done at the National Foreign Language Center in helping students to analyze their own performance in conversational interaction in Japanese, comparing this with the expectations of native speakers. This approach can be extended to studies of other individual language-learning and language-using strategies.

Computer data can enlighten us to the effectiveness of testing and, through its records and ability to vary content of tests, bring about genuine prochievement (that is, combined proficiency and achievement) testing (Clark's term, 1988:12). Through feedback and selective retesting, the computer can make testing a genuine aid to student learning, rather than a punitive 'Gotcher' approach.

The computer, laserdisc, and compact disc (CD-I audio and CD-ROM, see Woodbury 1988) are settling in as permanent features of society; it is for us to seek to use them effectively and economically to achieve our ends, rather than allowing them to determine our ends by their means. This will need to be done quickly, before what we do not want becomes set in silicon. 'To a great degree, the computer has been used most often to do what is already done, instead of for activities that have not yet been done because they *require* a computer' (Ariew and Frommer 1987:185). Our attention should be directed to advance in this area, while we increase and improve our own interactive teaching. Much effort and work lie ahead for many teachers and materials developers to ensure that expectations are realized and opportunities fully exploited. We would do well at this stage to heed Cervantes' advice: 'Patience and shuffle the cards.'

Appendix: Ten principles of interactive language learning and teaching.

1. The student is the language learner.

Corollary 1. Motivation springs from within; it can be sparked but not imposed from without.

2. Language learning and teaching are shaped by student needs and objectives in particular circumstances.

Corollary 2. Language teaching and course design will be very diverse.

3. Language learning and teaching are based on normal uses of language, with communication of meanings in oral or written form basic to all strategies and techniques.

4. Classroom relations reflect mutual liking and respect, allowing for both teacher personality and student personality in a nonthreatening atmosphere of cooperative learning.

5. Basic to language use are knowledge of language and control of language.

6. Development of language control proceeds through creativity, which is nurtured by interactive, participatory activities.

7. Every possible medium and modality are used to aid learning.

8. Testing is an aid to learning.

9. Language learning is penetrating another culture: students learn to operate harmoniously within it or in contact with it.

10. The real world extends beyond the classroom walls; language learning takes place in and out of the classroom.

References

- Ariew, Robert, and Judith G. Frommer. 1987. Interaction in the computer age. In: Rivers, ed. (1987:177-93).

- Clark, John L.D. 1988. Toward a research and development strategy for computer-assisted language learning. *CALICO Journal* 5.3.5-23.
- Davies, Graham D. 1989. CALL and NCCALL in the United Kingdom: Past, present, and future. In Smith, ed. (1989:161-80).
- De Francis, John, ed. 1951. Georgetown University Round Table on Languages and Linguistics 1951. *Linguistics and language teaching*. Monograph no. 1. Washington, D.C.: Georgetown University Press.
- Gleason, Jean Berko, ed. 1988. You can take it with you: Helping students maintain foreign language skills beyond the classroom. Center for Applied Linguistics series: *Language in education, theory, and practice*. Englewood Cliffs, N.J.: Prentice Hall Regents.
- Higgins, John, and Tim Johns. 1984. Computers in language learning. Reading, Mass: Addison-Wesley Publishing Company.
- Hope, Geoffrey R., Heimy F. Taylor, and James P. Pusack. 1984. Using computers in teaching foreign languages. Center for Applied Linguistics series: *Language in education. Theory and practice*. Orlando, Fla.: Harcourt Brace Jovanovich.
- Leech, Geoffrey, and Christopher N. Candlin, eds. 1986. *Computers in English language teaching and research*. London: Longman.
- Mueller, Marlies, Garth M. McCavana, Maureen Ramsden, and Sharon Shelly. 1985. Language learning laboratories: The end of a lukewarm affair? *French Review* 60.5 (1987):585-91.
- Murray, Janet H., Douglas Morgenstern, and Gilberte Furstenberg. 1989. The Athena language-learning project: Design issues for the next generation of computer-based language-learning tools. In Smith, ed. (1989:97-118).
- Otto, Frank. 1988. Using the computer. In: Gleason, ed. (1988:71-91).
- Phillips, Martin. 1986. CALL in its educational context. In: Leech and Candlin, eds. (1986:2-10).
- Rivers, Wilga M. 1987. *Interactive language teaching*. Cambridge: Cambridge University Press.
- Rivers, Wilga M. 1981. *Teaching foreign-language skills*. 2d ed. Chicago: University of Chicago Press.
- Rivers, Wilga M. 1989. Ten principles of interactive language learning and teaching. Prepublication paper. Washington, D.C.: National Foreign Language Center.
- Robinson, G.L. 1989. The CLCCS CALL study: Methods, error feedback, attitudes and achievement. In: Smith, ed. (1989:119-34).
- Robinson, Gail L., John Underwood, Wilga Rivers, José Hernandez, Carollyn Rudesill, and Claire M. Enseñat. 1985. Computer-assisted instruction in foreign language education. San Francisco: Center for Language and Crosscultural Skills. EDRS ED 262 626.
- Smith, Wm. Flint, ed. 1989. *Modern technology in foreign language education*. Lincolnwood, Ill.: National Textbook Company.
- Verano, Miguel. 1989. USAFA interactive study in Spanish. In: Smith, ed. (1989:249-56).
- Woodbury, Verl. 1988. CD-ROM potential and practicalities. *CALICO Journal* 6.1.25-35.
- Young, Richard. 1988. Computer-assisted language learning conversations: Negotiating an outcome. *CALICO Journal* 5.3.65-83.

Technological, methodological, and assessment challenges: Can the foreign language teacher survive?

Ray T. Clifford

Defense Language Institute

New technological, instructional, and testing innovations abound in the field of foreign language teaching today, but these are not threats to teachers. The foreign language teaching profession has survived the advent of many innovations and changes--some of which have had a lasting impact and many more which were short-lived. However, while the profession has learned to copy with pendulum-like swings in preferred teaching and testing approaches, it has not done as well in identifying and addressing the systemic problems which plague the field.

The following historical fable is offered as a reminder of the many solutions offered over the last three decades as solutions to a problem which is larger than any particular instructional strategy or technique.

The foreign language schoolhouse. Once upon a time there was a royal kingdom which had a little red schoolhouse. The little red schoolhouse was dedicated to the teaching of foreign languages. It was a lovely schoolhouse. And since the King had decreed that all students had to pass through that schoolhouse as part of their educational rites, the teachers were very happy. It was an academic paradise. The schoolhouse smiled.

Then one day the students became restless. 'Our training is not relevant,' they chanted in unrelenting pattern drills. To make matters worse, the King's court went on an around-the-world tour to study the problem. When they returned they joined the rising chorus of discord. 'Your students can't hack it,' they said in their refined diplomatic language and they dropped foreign language requirements. The schoolhouse was befuddled. Hadn't it always done its best? It longed for peace and tranquillity, for the halcyon days of required foreign language study.

Meanwhile, news of the ruckus reached the King, who sent his wizards and wise men to solve the problem. Each visitor arrived at the schoolhouse in an elegantly decorated horse-drawn wagon and, since they were the King's representatives, their arrival was heralded by a brass band. However, this pomp and ceremony soon became old, and the school, instead of extending a friendly welcome to these visitors, began to say, 'Oh, no! Not another bandwagon!'

One by one the King's experts continued to arrive. One visitor proposed, 'We should use a grammar translation methodology,' but the teachers responded that that method had lost its gloss. Another suggested the

audiolingual method with its motto 'parrot or perish.' But that method didn't succeed because it missed its cue. Even the new key methodology failed to unlock the door to foreign language learning. And although some tried, neither teachers nor students could break the cognitive code.

So wizards and wise men, sages and seers continued to visit. One capital idea was to finance improvements through federal funding, but even that failed to earn the public's long-term interest. Criterion-referenced instruction might have been the answer, but after developing thousands of objectives, teachers decided that objectives were objectionable. Some liked the microwave teaching method; others decided it was a half-baked idea. The direct method might have been chosen as the only true method, but even the most ardent proponents of the method couldn't explain the meaning of the target language word for 'truth' without resorting to English. One visitor suggested that we should adopt the eclectic method, but that idea was rejected because teachers (who of course are language experts) would not use a definite article with the word 'eclectic.' Even the visitor pushing communicative competence failed to get her idea across. Then a wizard with a cuisinaire rod asked for volunteers to use the silent way method, but no one spoke up. About this time another visitor arrived, playing baroque music and suggesting after all of the other suggestions, suggestopedia. His slogan 'Go for baroque!' was catchy, but the students were too comfortable to follow up on it. From another direction came a complete physical response methodology, which worked until they tried to teach the phrase, 'Jump out the window.' Unfortunately, it was a two-story schoolhouse. Still, no one was extremely concerned, because all language teachers were experiencing falling enrollments.

The little red schoolhouse became anxious. Without students, would it be torn down? About this time testing experts visited the schoolhouse with their solutions. At first the discrete point testers had all the answers. Then the pragmatists came with the claim that if teachers were to find the naked truth about their students' language proficiency, they had to use cloze tests. This apparent contradiction just left everyone feeling testy. Even the textbook writers decided to turn a new leaf. A European curriculum developer soon arrived, but the teachers didn't understand his function and had no notion of where to begin. Computer experts brought in innovative programs, but they were not accepted. Some teachers feared that computer assisted language study would produce terminal learners--others, that it would lead to over-monitoring.

The little red schoolhouse was despondent. It wanted to cry. Its windows misted up. Retreating into a state of contemplation, it asked itself, 'What is the problem? We all agree on the goal of our instruction. The students want to be fluent speakers of the language they are studying. The teachers want their students to be fluent speakers. The King's court and the major employers in the country also want the students to be fluent speakers of the languages they are studying.'

'Why have we not found the solution?' moaned the schoolhouse in despair. Sadly, it never occurred to the little red schoolhouse that no one, not even the language experts, had defined what 'fluent' really meant.

There are two morals to this story. The first is the oft-repeated, 'If you don't know where you're going, you won't know when you've arrived.' The other is, 'If you don't know where you're going, you have an unlimited number of options for getting there.' When applied to a professional discipline such as foreign language teaching, either conclusion is disconcerting. The fact that this condition has existed for so long is deplorable.

Although we may be reluctant to admit it, foreign languages have not been very visible in our nation's history. The *Almanac of American History* has no entries about foreign languages. The *Almanac* does report that in July of 1789, the year Georgetown University was founded, the first Congress moved into session. Shortly after convening, it created the first executive department for the new United States government, the Department of Foreign Affairs. But this visibility for 'foreign' activities disappeared less than two months later when Congress changed the department's name to the Department of State.

The history of foreign language teaching from that point forward has been one of continuing struggle. At first the struggle was against other languages. It took nearly a century for modern languages to replace ancient Greek and Latin in our accepted academic curriculum. That was finally accomplished, but at a cost. To prove that modern languages were educationally challenging, we accepted a detached analytical approach to teaching, a two-year language sequence and limited instructional objectives. It seems that in order to convince the educational establishment that modern languages were as intellectually challenging as classical Greek and Latin, we had to make the learning of these new courses as difficult as possible. The more difficult the language learning process appeared, the more credibility the profession enjoyed. The broadly accepted objective of foreign language study at the time was clearly not to build communicative skills, but to discipline the intellect.

That led, even a hundred years ago, to some criticism. The following quote is from the preface of a textbook published in the year 1882.

In our times, when international commerce and intercourse is so costly increasing, our schools and colleges must aim at other and more practical results than heretofore were considered necessary.

It is not longer sufficient to teach the student the grammatical peculiarities of French and German, and to introduce him into the classic literature of these languages; but the true end and aim of our linguistic education must be to actually speak the modern tongues, and to really be able to converse in them fluently and idiomatically.

'The usual mistake,' says the *New York World*, in an able editorial on the study of modern languages, 'in America, throughout the majority of schools is that in studying a foreign tongue more actual study is put upon English and a formation of a smooth translation than in building up and acquiring the language in question. But whatever the faults of teachers or of the system, of one fact the parents and public are painfully assured, and that is after years of study the scholars are still unable to speak and write the language and with difficulty can even read it' (Rosenthal 1882).

Although I'm sure these words resonate with a majority of today's foreign language teachers, there is no evidence that they had any impact on the profession at the time.

By the beginning of the twentieth century modern languages had firmly established themselves as part of the public school curriculum, but classical teaching objectives and methods persisted. William Riley Parker reports that in 1915, almost as many high school students were studying modern languages (35.9%) as were studying Latin (37.3%) (Parker 1961:85).

Then World War I brought about dramatic shifts in the teaching of foreign languages in the United States. By 1922, the percentage of high school students studying German had dropped from a high of 24.4 percent to 0.6 percent. Spanish and French students had in part filled in the gap, but language study had received a setback. In fact, in that time frame--during and shortly after World War I--22 states passed laws which restricted the teaching of foreign languages. In 1923, the Supreme Court overruled those laws, but an antiforeign and antiforeign language mindset was still apparent (Parker p. 85).

No sooner had this Supreme Court ruling opened the door, than the Carnegie-supported Modern Foreign Language Study closed it. The most influential portion of that study, the Coleman Report, found--to no one's amazement--that 83 percent of the students studying foreign languages in high school completed only two years of study. It also noted that colleges accepted those two years as sufficient to fulfill their foreign language entrance requirements. The 'natural' conclusion drawn was that if only two years of study time were available, the schools should restrict themselves to the teaching of reading, which the report reasoned might be accomplished in that amount of time.

Thus, the emphasis on reading and translation continued, and while foreign languages remained in the schools, they were effectively excluded from the high school 'core' curriculum. Parker (p. 89) mentions three reports from the early 1940s which reinforced this secondary status of foreign languages: *What the High Schools Ought to Teach*, published in 1940; *Education for All American Youth*, published in 1944; and *General Education in a Free Society*, published in 1945. All of these left foreign languages out of their core curriculum proposals for high schools, because such courses took time which could more profitably be devoted to other subjects.

When World War II arrived, our nation found itself without people who could communicate in critical foreign languages. This realization could have been a boon for the foreign language profession, but it was not. On the positive side, the 'Army method' of language teaching, which was created to meet critical wartime needs, proved that given enough time, with select students, we could teach people to communicate in foreign languages.

Unfortunately, the total study time required was not emphasized, and all the general public noticed was that the Army was somehow teaching foreign languages in nine months. Even within the profession, few noted that those nine months included the class-hour equivalent of six years of foreign language study. As a result, the profession at large suffered from the perception that if the right method were used, public schools could also teach a language in nine months.

Despite this naive attitude, our awareness of the need for foreign languages grew and enrollments slowly continued to climb until the early 1950s. Then Sputnik circled the globe. The National Defense Education Act of 1958 brought unprecedented publicity, increased enrollments, teacher workshops, federal funding, and increased respect for the profession. It was at this time that William Riley Parker published *The National Interest in Foreign Languages*. In looking at this postwar period Parker reports (1:92):

In a few universities, e.g., Cornell, the 'Army method' was hopefully adapted to academic teaching; in some institutions language instructors hopefully put more emphasis on speaking skills; 'language laboratories' were enthusiastically installed in a number of places. In most places, however, foreign language teachers settled back with a sigh to cope anew with their ancient formidable enemy--lack of time. *Nowhere in educational circles did anyone in authority propose that we now follow European example, or learn the main lesson taught by the 'Army method,' and allow sufficient time for language study to make its results meaningful.*

Despite this somber recognition of reality, Riley's view of the future was unabashedly optimistic. From the perspective of improvements that had already been attained under the NDEA, Parker ends his seminal work with some predictions about the future. One of those was, 'Eight and ten year sequences of foreign language study will be common in the public schools.' This has clearly not happened. He also predicted that 'The better colleges and universities will require demonstrated proficiency (not high school "units") in a foreign language for entrance, and demonstrated proficiency in a second foreign language (often non-Western) for graduation' (p. 153).

Sadly, we have not made the progress Parker predicted, and our lack of progress is, I believe, directly attributable to a lack of ability to define progress. If Parker's latter prediction had materialized, the former would have come to pass as well. The first step in making progress is to define the goal, and the only way we will be able to ascertain whether we are improving is to be able to measure it.

The foreign language teaching profession was heartened by the recommendations of the Commission on Excellence in Education (Gardner, 1983). However, the reality of our present situation is that compared to other countries, foreign language instruction in the United States has yet to achieve mediocrity. Unfortunately, the recommendations in *A Nation at Risk* are based again on seat time and not demonstrated proficiency.

Without an ability to measure success, we will lack the credibility needed to follow through on the Commission's recommendations. Sadly, we cannot yet prove to administrators what we as teachers all know: The acquisition of foreign language skills requires more time than the educational system currently allows.

Although the terminology used by administrators is usually more erudite, success in academic programs is invariably defined by one of the three criteria: better, faster or cheaper. Unfortunately, budgetary pressures often force decision makers to ignore the 'better' option in favor of 'faster' or 'cheaper.' This danger is especially real when there is no working definition of 'better.'

In this regard the Defense Language Institute, in Monterey, California, was no different than the rest of the profession (Clifford 1987). From the mid-1960s until the early 1980s, student end-of-course grades were assigned on a norm-referenced basis. By the early 1980s, the disadvantages of this norm-referenced definition of excellence had become evident. No matter what program cuts were made, 50 percent of our graduates were above average. Without a criterion-based definition of 'better,' we were at the mercy of 'faster and cheaper' initiatives.

Faculty staffing ratios, for instance, declined to the point where the teaching load was becoming oppressive. Teachers were expected to teach so many hours that there was inadequate time available for preparation.

Eventually, faculty staffing ratios were reduced to the point that providing substitute teachers to cover unscheduled teacher absences began to be a real problem. Yet pressures for continued efficiencies within the academic program continued. To achieve maximum efficiency, teachers schedules were changed on a weekly and sometimes daily basis. As a result, teachers lost the continuity necessary to know their students and adapt classroom instruction to meet their students' needs. As teachers struggled under these adverse conditions, faculty morale suffered.

The situation was only reversed because the institute accepted the set of hierarchical proficiency standards developed by the Interagency Language Roundtable and then produced proficiency tests based on those standards. These proficiency standards and tests were implemented in 1985, and the student data from that year were considered a baseline against which further program enhancements are evaluated. Also in 1985, language needs assessments were conducted for the major occupations of our graduates, using the same proficiency standards. A comparison of the foundation language skills needed to do those jobs with the skills then attained by our graduates revealed that less than a third of our graduates possessed the level of language proficiency required in their future jobs.

The external recognition of this shortcoming in the quality of our instructional program provided the opportunity to address the problem of inadequate staffing, and a 30 percent increase in staffing was proposed as the central feature of our improvement plan.

As can be seen from Figure 1, the institute has improved its average teacher-to-section ratio from 1.5 in 1985 to 1.9 in 1988. During the same period, the percentage of students attaining the minimum proficiency levels identified in the 1985 language needs assessment has also increased (see Figure 2).

These results not only indicated that our improvement plan was working, they have added an important dimension to recent program budget decisions. As the institute was asked to justify its budget in light of pending Department of Defense budget reductions, these data provided crucial evidence that increased spending can result in increased learner outcomes. In fact, correlating the average staffing ratios with the average percentage of students attaining the required foundation skills for the last four years yields a phenomenal $r = 0.997$. Because correlations between average data points are somewhat inflated because of the reduced variance inherent in average scores, an even more persuasive statistic is that a 28 percent increase in staffing has

resulted in a 54 percent increase in the percentage of students meeting the desired proficiency standards.

Figure 1. Average instructor-to-section ratio.

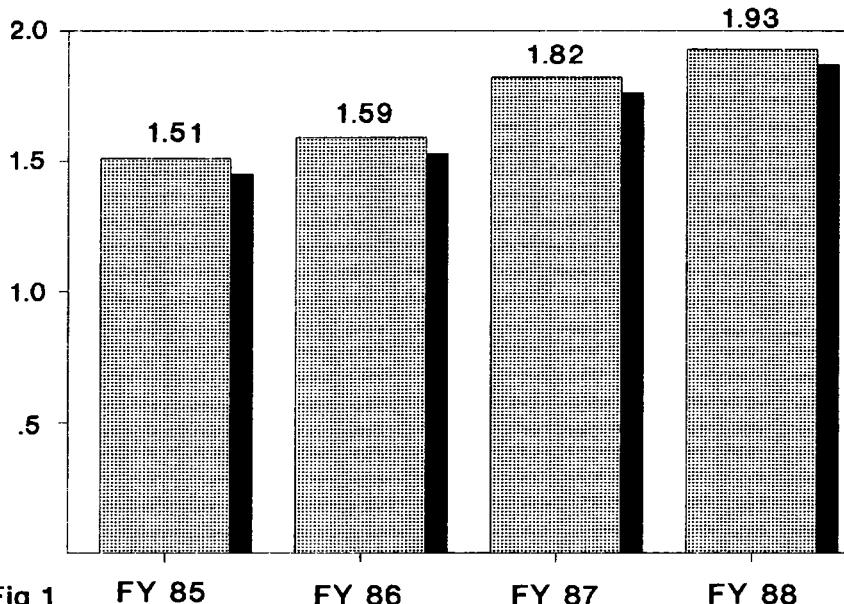


Fig 1 **FY 85** **FY 86** **FY 87** **FY 88**

Of course, it was not increased staffing alone which brought about these results. Increased staffing simply provided the opportunities for remedial instruction, better lesson preparation, and faculty in-service training which were not possible with more 'efficient' staffing ratios.

DLI's experience indicates that excellence is only attainable when it is defined and the best way to protect quality is to find a way to measure it. The foreign language profession might benefit from a similar approach. As we work to ensure that our future is not merely a repeat of the past, our ability to define success must be a high priority.

Figure 2. Students completing basic courses who met current graduation standard.

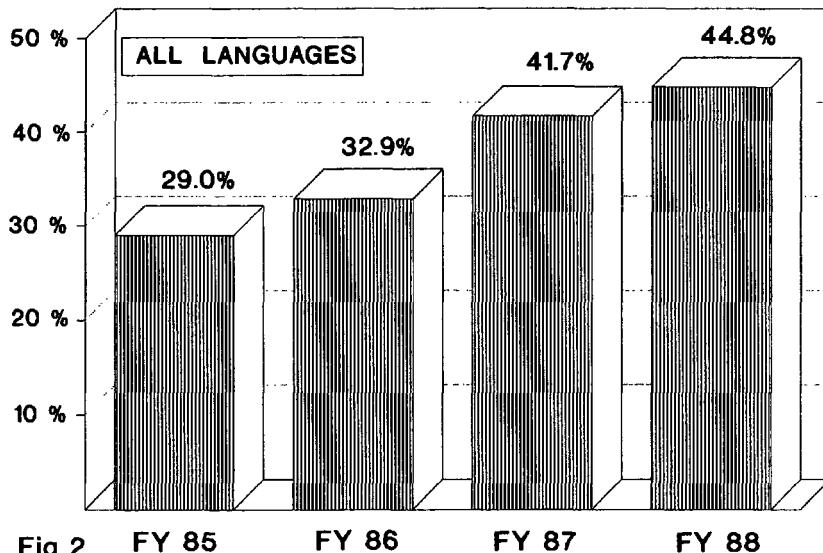


Fig 2 FY 85 FY 86 FY 87 FY 88

References

- Clifford, Ray T. 1987. Language teaching in the federal government: A personal perspective. *The Annals of the American Academy of Political and Social Science.* 490 (March 1987): 137-46.
- Gardner, David P., chair. 1983. *A nation at risk: The imperative for educational reform.* The National Commission on Excellence in Education. Washington, D.C.: U.S. Department of Education.
- Parker, William Riley. 1961 (released 1962). *The national interest and foreign languages.* 3rd ed. Washington, D.C.: Department of State Publication 7324, U.S. Government Printing Office.
- Rosenthal, Richard S. 1882. *The Meisterschaft system.* Boston, Mass: Meistershaft Publishing Company.
- Schlesinger, Arthur M., Jr., ed. 1986. *The almanac of American history.* New York: Bramhall House.

Multipurpose language tests: Is a conceptual and operational synthesis possible?

John L. D. Clark
Defense Language Institute

The official title of this paper is very much in keeping with the traditions and general manner of operation of the Georgetown University Round Table on Languages and Linguistics: staid, proper, sophisticated, erudite. This year, however, through the combined efforts of several Round Table speakers, including keynoters Protase Woodford, Stephen Krashen, and Ray Clifford, the discourse-level barricades have been breached, permitting the introduction of a soupçon of humor. I accordingly hasten to proffer an alternative title: 'Can There be Shmoos in Language Testing?' For the information of those in the audience who happen to postdate the belt-tightening days of World War II, Shmoos were the fanciful animals created by cartoonist Al Capp and prominently featured in his *Li'l Abner* comic strip in a valiant and reasonably successful attempt to help America make light of its then-current woes.

Roughly the size and shape of pregnant bowling pins, Shmoos cost nothing and required no upkeep. Some Shmoo owners were of the opinion that cooked Shmoo tasted just like ham; others found the taste very similar to that of chicken. In any event, Shmoos loved to be killed and eaten, and those that did not go straight into the oven or frying pan spent their leisure time having a great number of progeny, all equally tasty. Shmoo skin made wonderfully flexible and durable shoes, and various other Shmoo parts served a number of other important and practical functions. In short, Al Capp's cartoon creation was, in a very significant sense, all things to all people.

Can there also be Shmoos in language testing? Can one single type of instrument flexibly serve a number of differing measurement goals and purposes? In addressing this question, it will be helpful to make use of the six short guideline questions made famous by the newspaper guild: Who?, What?, When?, Where?, How?, and Why?

Answers to the *Who?* and *Where?* questions with respect to language testing are relatively straightforward--the *Who?* being the examinee taking the test and the *Where?* the location of the testing event. It should be noted that the term 'examinee' is more appropriate than 'student,' because persons other than formal students can and do take language tests in noninstructional contexts, for example, in connection with employment application or, in some instances, simply for intellectual or self-evaluative purposes. By the same token, although the *Where?* is most typically the foreign language classroom, possible testing locations can and do include numerous other sites, such as large auditoriums (e.g. for college entrance examinations), company offices, and even quonset huts in Third World villages (e.g. for in-country testing of Peace Corps volunteers).

The *When?* of language testing (i.e., the specific occasions on which language tests, of whatever type, are administered to examinees) is an interesting issue that could readily justify separate treatment beyond the scope of the current discussion. For present purposes, I would simply suggest the wide-scale adoption of a 'zero-based budget' approach to test administration, in which language testing would never be conducted simply for its own sake, but only as and when necessary to accomplish demonstrably important educational or other societal purposes. We need continually to examine our testing activities to insure that some pragmatically significant use is being made of the results. For example, in our language classes, do we give a mid-term exam simply because there was a mid-term last year and the year before that? Or because we always had a mid-term exam when we ourselves were language students? A strict application of 'zero-based budget' (as well as 'sunset-law') principles to existing or proposed language testing activities would considerably sharpen our operational focus as to the nature, necessity, and appropriate frequency of these activities.

The three remaining guideline questions are *Why?*, *What?*, and *How?* The *Why?* is concerned with the rationale, purpose, or reason for which the testing activity is being carried out in the first place. The *What?* and *How?* together address the 'mechanics' of test development and implementation in the service of a particular, previously defined and elaborated testing 'purpose'. It is not possible to overstress the proper sequence in which these questions should be addressed: the *Why?* must always come before the *What?* and *How?*. It would not be overstating the case to say that if the test developer is scrupulous about elucidating the *Why?* as the first order of test development business, the *What?* and *How?* will virtually take care of themselves, in the sense that the appropriate test content and procedure will in large part be implicit in, or straightforwardly derivable from, the test's intended purpose. By contrast, if *What?* and *How?* considerations are emphasized too early in the overall development process, the test designer runs a substantial risk of losing track of, or doing conceptual and operational violence to, the informational goals the test was ostensibly intended to meet.

This brings us to the sequentially last--but operationally first and most important--question: *Why* are we testing? Notwithstanding its critical significance, the *Why?* question is very easily resolved by virtue of the fact that, at least at the basic conceptual level, there are only two possible reasons or purposes for giving a language test: (1) to provide 'instruction-oriented feedback' to teachers and/or students within a pedagogical framework; or (2) to provide information relevant to 'real-life decision making' outside of or separate from language learning/instructional issues per se.

It is, I would suggest, unfortunately but unavoidably the case that no single testing instrument--regardless of the care, diligence, and insight with which it is developed--can optimally serve both instruction-oriented feedback and real-life decision-making purposes. The situation is analogous to that involved in the design of 'flying boats' and other multipurpose contrivances often featured in publications such as *Mechanix Illustrated*. Although the typical flying boat can and does manage to propel itself through both water and air, the design compromises that are involved result in a vehicle that falls considerably short of performing in either element with the suitability,

appropriateness, and overall usefulness of devices designed and built explicitly for use in only one of these environments. The analogous problem facing foreign language test developers is to make an informed and well-reasoned choice between bending all test development efforts toward accomplishing one of the two possible testing purposes (instruction-oriented feedback or real-life decision making) or attempting to develop a single test that simultaneously addresses both purposes, but to a necessarily less than optimal extent in both instances.

For purposes of discussion, let us explore the nature and consequences of adopting a single-purpose testing orientation, examining first the design and development issues associated with tests that are frankly and exclusively aimed at providing 'instruction-oriented feedback.' In this regard there is, I believe, one primary touchstone or development principle that must be assiduously understood and followed. Specifically, the test must be designed in such a way that the feedback provided is of an appropriate type and degree of detail to permit straightforward instructional remediation. But how much and what type of feedback is called for? Let us consider a potentially useful medical analogy, specifically, a visit to the doctor for diagnosis and treatment of a tired, ache-all-over, slight fever, stopped up nose, frequent coughing condition that has been going on for about a week. If, after a ten-minute (and sixty-dollar) examination and lab tests, the only feedback to be provided were in the form of the statement, 'You're sick,' this would not be particularly enlightening or adequately indicative of what specific steps might be taken to improve the situation. At the other extreme, a response along such lines as, 'Your serum albumin is elevated, your hemocrit level is down, and you have a slight electrolyte imbalance' would, although considerably more detailed than the preceding, also fall short of providing truly useful remedial information. 'You have a slight case of the flu' would, however, provide highly appropriate feedback, since this particular information could be used, either in conjunction with the patient's own prior knowledge of the proper treatment of flu conditions or with the added assistance of a doctor-provided instructional leaflet, to initiate a suitable and effective treatment program.

In the language testing context, instruction-oriented feedback in the form of, for example, 'You're a level 1+' would provide virtually no information about areas of strength or weakness in the examinee's language performance or suggest suitable instructional interventions. 'Your apico-dental unvoiced plosives are weak in syllable-initial position' would, at the other extreme, err in providing excessive amounts of information in a form that would not be of practical utility to the typical (nonphonetics major) student. Testing feedback of the general form, 'Your knowledge of food and restaurant vocabulary is pretty good, but you need to work more on using the proper gender form of adjectives' would be more immediately meaningful to the student and would more clearly indicate the additional learning tasks at issue.

An important corollary to the need to provide adequately detailed instruction-oriented feedback via the testing procedure itself is the need to insure that, on the instructional side of the equation, suitable learning material and opportunities are available to the student to address and help correct the various deficiencies identified. In keeping with the previously proposed zero-based orientation to test development and use, it would make both

conceptual and practical sense to suggest (in the instructional-feedback testing context) the additional guideline principle, 'Don't test what you can't teach.' To develop and administer a highly diagnostic feedback test without at the same time providing students the opportunity and wherewithal to correct the identified deficiencies would be akin to presenting a hungry diner a sumptuous menu in a restaurant that has no food.

One example of a very thoroughgoing approach to the design and development of explicitly single-purpose diagnostic feedback tests is the series of test batteries developed under a recent joint project of the University of Pennsylvania and the Center for Applied Linguistics to study the nature, and relative rates, of loss of various second language skills following cessation of formal instruction (Lambert and Stansfield 1988). Under this project, tests of 'listening lexicon,' 'listening structure,' 'reading lexicon,' 'reading structure,' 'speaking lexicon,' and 'speaking structure' were developed for three different languages: Arabic, Chinese, and Japanese. As suggested by their respective titles, the 'listening lexicon' test was intended to determine the examinee's understanding of individual lexical items in an aural comprehension context and the 'listening structure' test was designed to check the comprehension of spoken grammatical structures. 'Reading lexicon' and 'reading structure' were aimed at determining the examinee's comprehension of individual lexical and structural items presented in written form; and the 'speaking lexicon' and 'speaking structure' tests required the examinee to produce target language lexical and structural items orally. To eliminate or minimize the influence of extraneous variables on examinee performance, a number of specialized procedures were followed. For example, to avoid the problem of chance success associated with multiple-choice testing, the project tests were all of the 'constructed response' type, in which the examinee was required to write out (or, in the case of the 'speaking lexicon' and 'speaking structure' tests, say aloud) the intended response rather than choose from available options. Specifically, the stimulus-response conditions for each of the tests in the battery were as follows:

	Stimulus	Response
Listening Lexicon	Spoken TL sentence	Written NL translation
Listening Structure	Spoken TL sentence	Written NL translation
Reading Lexicon	Printed TL sentence	Written NL translation (of underlined lexical item)
Reading Structure	Printed TL sentence	Written NL translation
Speaking Lexicon	Printed NL sentence	Spoken FL translation
Speaking Structure	Printed NL sentence	Spoken FL translation

For each of the tests in the battery, the stimulus sentences were deliberately kept as short as possible (consisting, in many cases, of only four or five words). For the 'lexicon' tests, the structures in which the lexical items were embedded were as simple and straightforward as possible. Correspondingly, in the 'structure' tests, the lexicon used was uniformly extremely simple and presumably well known by any examinee taking the test. Test scoring was on a right-wrong basis for each item, taking into account only the specific linguistic element at issue. For example, in the structure tests, the response

was considered correct if the examinee properly translated (for listening/reading) or produced (for speaking) the structural feature being tested, notwithstanding any possible mistranslations or omissions of the lexical items in which the structures were embedded.

Detailed item-level results of project test administrations indicated quite clearly that the highly discrete, targeted testing approach used in the Pennsylvania-CAL study was capable of providing extremely detailed information as to the examinee's control or lack of control of the particular lexical and structural features included in these instruments, with this diagnostic capability extending also to the particular skill modality (listening, reading, speaking) at issue in each instance. However, this high degree of analytic precision was necessarily obtained at the expense of reduced potential applicability to other more generalized language-use contexts. From both face- and content-validity standpoints, the examinee's ability to, for example, comprehend a spoken lexical item in the context of an intentionally very short and simplified utterance would not ipso facto be considered presumptive of a similar ability in more integrated, naturalistic settings typical of real-life language use.

To look briefly at the prospects for diagnostically oriented feedback testing over the near- and longer-term future, I would suggest that this type of testing represents a highly appropriate application for computer-assisted language learning (CALL) projects. Within the CALL setting, it would be readily possible to prepare computerized testing programs incorporating the detailed, highly specific testing formats and exercises that are needed to yield fine-grained diagnostic information and, at the same time, make available to the student, also via appropriate programming, the particular targeted learning activities that would be most efficient and effective in remediating the identified learning shortfalls. Delegating all or most of the required 'discrete item' diagnostic testing and remediation tasks to the computer would, as an additional benefit, allow the teacher more time and opportunity to work with the students in more highly authentic discourse settings (e.g., free or semi-structured classroom conversation) in which the teacher would be able to informally note and communicate useful information to the student about his or her performance in these more naturalistic contexts.

Let us turn now to the second of the single-purpose testing approaches: the development of tests aimed exclusively at 'real-life decision making.' By way of example, it would be useful to consider how one might best go about developing a test to serve the real-life decision-making purpose of selecting individuals to monitor foreign radio broadcasts on an around-the-clock basis. A test designed explicitly for this purpose would probably attempt to reproduce, to the greatest extent possible, the particular types and conditions of language use that would be encountered in the real-life situation at issue. Broadcasts of a variety of types and content representative of those present in real-life transmissions would be obtained or produced, and the quality of the recordings would be adjusted so as to reflect as closely as possible the signal-noise ratios, types and extent of static, and other technical characteristics of the real-life transmission. The amount of time available to the listener during the test to receive and react to the aural stimuli would deliberately be set to correspond to the time actually available in the real-life

situation, and so forth. The probable end result of the development effort for such a highly realistic, real-life reflecting testing procedure would be an assessment instrument that on face-, content-, and operational validity grounds could be viewed as closely and appropriately reflective of probable actual competence for the particular real-life task at issue.

Unfortunately, tests designed to maximize accuracy of prediction for a particular specialized real-life language use task are, by the same token, generally less capable of predicting performance in areas other than the targeted application. In the previously cited radio broadcast monitoring situation, for example, although the quite high level of specialization incorporated into both the content and manner of operation of the testing procedure serves to enhance the utility and validity of the test for its specific intended function, this advantage is gained at the expense of potentially reduced applicability to other types of listening comprehension situations not adequately represented or approximated in the content or procedures of that particular test (in this instance, for example, aural comprehension within dyadic conversations in which the examinee functions as a real-time interlocutor rather than a passive listener).

Much of the science or art of developing tests aimed at real-life decision making involves a principled decision as to where a particular test should fall along the bipolar continuum of predictive accuracy vs. range of extrapolation. A good example of a pragmatically useful compromise between these two extremes is the series of testing instruments developed and used in the Graduate School Foreign Language Tests (GSFLT) program of the College Board (Clark 1968; Harvey 1968). The real-life decision-making purpose of the GSFLT is to predict the examinee's ability to read journal articles and other texts typical of those encountered by scholars in a wide number of graduate-level specialization fields. Although the theoretically optimum approach to predicting, for example, the examinee's ability to read with comprehension texts on nuclear physics would be to present a variety of textual excerpts drawn directly from this particular area of specialization, the extremely large number of areas that would have to be assessed, by means of similarly discrete instruments, in order to adequately cover the range of specializations available within the American graduate education field made this approach clearly unfeasible. On the other hand, a test consisting solely of generic, general-purpose reading passages was considered not to have adequate face- or content-validity as a measure of the examinee's ability to deal with specialized textual material.

The compromise solution ultimately adopted by the GSFLT program was to develop a test which included--in addition to a 'core' test section on basic structures and lexicon--three separate sections of reading materials drawn from humanities, natural sciences, and social sciences sources. The examinee was instructed to select and work with the one of the three sections that most closely approximated his or her area of specialization. Under this testing procedure, although the nuclear physics student would not be asked to read passages drawn directly from that particular field, there would nonetheless be the opportunity to demonstrate competence in working with passages drawn from the natural sciences area, which would be expected to approximate more closely the style and content of 'nuclear physics' discourse than would more

generalized, undifferentiated texts and, as such, be more highly predictive of the specialized ability.

The Interagency Language Roundtable (ILR) verbal descriptions of language proficiency and associated testing techniques, derived from earlier descriptions and procedures developed by the Foreign Service Institute (FSI) of the U.S. Department of State (see Wilds 1975; Clark and Clifford 1988), also reflect an implicit if not formally delineated attempt to strike a pragmatically useful balance between predictive accuracy and range of extrapolation. In this regard, the verbal descriptions associated with each individual ILR level are intended to be broadly predictive of the examinee's performance in real-life language use situations similar to those referenced in the verbal descriptions. For example, in the case of an examinee scoring at level 3 on the ILR scale, the strong inference is made that he or she has a solid spoken command of the basic structural features of the language, has a pronunciation clearly comprehensible to native speakers not accustomed to dealing with foreigners, and a vocabulary adequate for talking fluently, albeit with occasional circumlocution, about a variety of practical, professional, and social topics. Although more highly focused performance tests are unarguably better able to assess functional competence for specified, delimited applications, the intentionally more generalized nature of the ILR scale and testing process permits extrapolation of the obtained results to a wider variety of contexts. Alternatively stated, the ILR approach may be viewed as trading off extreme accuracy of prediction for a particular language-use situation in favor of a reasonably good estimate of potential performance in a greater variety of situations.

Let us now turn to the final question before us today: the possibility of simultaneously accommodating the two separate purposes of 'instruction-oriented feedback' and 'real-life decision making' within a single test instrument. A reasonably detailed discussion of two somewhat different attempts at implementing this dual-purpose testing approach--'supplemented' FSI/ILR testing procedures, and the recently introduced concept of 'prochievement' testing--will exemplify some of the theoretical and practical considerations involved.

Since the introduction of the FSI interview procedure and scale in the early 1950s, a number of attempts have been made to buttress the original 6-level (0, 1, 2, 3, 4, 5) rating scale with more detailed information about the examinee's linguistic strengths and weaknesses. Indeed, the 'plus' ratings (1+, 2+, etc.) that were added to the original scale shortly after its introduction reflected the desire to provide more fine-tuned feedback on examinee performance than was possible using only the original scale. This expanded scale did provide a limited degree of instructionally relevant feedback within known instructional contexts. For example, students rated at level 2+ could, in the majority of instances, be considered to need greater work in the area of formal grammar, since developed experience with the rating scale tended to indicate that the limiting element preventing the awarding of level 3 was, typically, insufficiently accurate control of the basic grammatical structures required for full performance at the '3' level. Course-of-interview notes on examinee performance taken by one of the two members of the interviewing team (and which dealt for the most part with structural, lexical, and phonetic

aspects) were also regularly discussed with the examinee in a debriefing session following the interview.

A more highly structured approach to providing instructionally useful information in the course of the FSI interviewing and rating process involved use of the so-called 'factor scores,' used both at FSI and, with some later adaptations, by Educational Testing Service, the Defense Language Institute, and other government language teaching organizations (Educational Testing Service 1979). Under this approach, in addition to receiving a single global proficiency rating at the end of the interview, the examinee was evaluated on his or her performance in each of the five major linguistic performance categories or 'factors' at issue in the global rating scale: pronunciation, grammar, vocabulary, fluency, and listening comprehension. For each of these factors, a five-cell (representing levels 1 through 5) scoring grid was developed, with each cell in the grid containing a brief verbal description of the type and quality of performance on that factor that would be expected at that FSI level. For example, the verbal description associated with level 2 performance on the 'grammar' factor was 'fair control of most basic syntactic patterns; conveys meaning accurately in simple sentences most of the time.' To reach level 3 with respect to the 'vocabulary' factor, the examinee's lexical control would have to be 'adequate for participation in all general conversation and for professional discussions in a special field.' By circling or otherwise indicating the particular level score assigned to each of the five broad linguistic factors, the examinee's 'performance profile' across these factors could be determined and some degree of instructionally relevant feedback provided as to which of these factors were adequately controlled and which required improvement in order to advance to a higher overall level on the global FSI scale.

Although these and other 'supplemented' FSI scale-based testing techniques may be considered to provide at least some instructionally useful feedback in addition to serving their major real-life predictive purpose, it should be emphasized that the procedures involved are considerably less discrete and less specifically focused than are those utilized in tests designed solely and explicitly for instructional feedback purposes (e.g., the Pennsylvania-CAL lexicon/structure tests previously described). As a result, the overall contribution of the former procedures toward the specification of detailed remedial activities on the part of the language learner is necessarily rather limited by comparison to single-purpose instructional feedback instruments.

The term *proachievement* testing was coined by the speaker in the course of a small-group working conference on achievement testing convened in 1982 at Educational Testing Service by the National Association of Self-Instructional Language Programs (NASILP). A major concern of this conference was the design of tests of language performance that could be used not only to determine the extent to which students in a given self-study program had mastered the particular linguistic items presented in that program's syllabus and textbook (i.e. instructionally oriented feedback testing), but also to provide some information on the extent to which the examinees were able to use the language for functional communicative purposes within a realistic language-use setting (real-life performance assessment). One

approach with considerable conceptual appeal, and that had in fact been tried out on earlier occasions by a number of NASILP testers, was to attempt to elicit actual communicative use of the language on the part of the examinee (for both reception and production), but to do so within the context--and associated limitations--of the particular lexicon and structures covered in the textbook materials as of the time of testing. This approach thus combined both 'proficiency' and 'achievement' assessment elements, with at least the theoretical possibility of providing information of some degree of utility for both assessment purposes.

It is fair to state that the 'prochievement' testing approach, although of considerable interest from both theoretical and practical perspectives, is at present in only the very early stages of development, tryout, and debate. One important issue is that of the degree of naturalness or authenticity that can realistically be incorporated into the testing situation, especially in the early stages of instruction, where very little linguistic material is available to the student in either the lexical or structural domains. In an earlier discussion of what she referred to as 'hybrid' tests, Omaggio (1980:435) suggested that tests based on an achievement-plus-proficiency principle would, among other things, need to be written in such a way as to insure that: 'the situation depicted is relevant and immediately useful to the language learner...the language is, at all times, natural, respecting the "conditions of elicitation" of certain types of structures in natural language use...answers required of students have "truth value"...[and] respect sociolinguistic norms'. To meet these criteria within the highly restricted language corpus available to the early-stage learner would be a substantial test design and delivery challenge.

Notwithstanding the early and understandably not yet fully formed nature of the 'prochievement' testing undertaking, tangible potential advantages to this dual-purpose approach can be discerned. Since the total scope of linguistic material presented to (or elicited from) the student is quite restricted, instruction-relevant feedback on his or her performance on given language aspects can be much more direct and precise than would be possible in the typical FSI/ILR test situation. At the same time, the operational approximation, within the testing process, of the informational requirements and elicitation conditions present in the real-world setting should have a very salutary washback effect on instruction, since this orientation would tend to guide teachers, students, and, ultimately, textbook writers and other instructional materials developers toward the intended outcome goal of genuine functional competence in the language. On the other hand, one caution to be raised in connection with 'prochievement' testing activities is the need to keep always in mind that the results of a test that is deliberately restricted to particular lexical domains or structural elements cannot legitimately be extrapolated beyond the bounds of the subject matter actually covered in that particular course and its associated test. The practical result of overlooking this caveat would probably be to overestimate the student's ability level by comparison to the more modest performance that would be demonstrated on an FSI/ILR-type proficiency test--that is, on an instrument designed to assess the examinee's overall performance across a wide range of language-use situations independently of the content of any particular text or curriculum.

What shall we take as the summary or 'bottom line' recommendations from our discussion today? Is a useful synthesis of testing purposes possible? Can there be Shmoos in language testing? My own inclination would be to suggest that the theory and technical practice of language testing is best advanced if the two purposes of instruction-oriented feedback testing and testing for real-life performance prediction are kept conceptually and operationally distinct, and if each type of test is developed and validated according to its own rationale and standards. In the course of following one of these two purposes explicitly, there may, in some instances, be 'targets of opportunity' that can be taken advantage of to address the other purpose to some extent, as is the case, for example, in the 'supplemented' FSI/ILR procedures. Nonetheless, as a general matter, I would suggest considerable caution in overtly attempting to accomplish both functions at once, on the grounds that what we will most probably wind up with will not be a testing Shmoo but a testing Frankenstein which, rather than well serving two masters, becomes its own master, to the potential detriment of both instruction-oriented and real-world decision-oriented testing. At the same time, however, we should probably admit that 'to experiment is human,' and jointly agree that, without feeling too guilty, we should be allowed to keep whetting our testing research and development appetites by continuing to think and hope that a plump and friendly Shmoo (with only the tiniest little forehead zipper) that tastes exactly like tuna casserole is just around the corner.

References

- Clark, John L. D. 1968. The graduate school foreign language requirement: A survey of testing practices and related topics. *Foreign Language Annals* 2.2:150-64.
- Clark, John L.D., and Ray T. Clifford. 1988. The FSI/ILR/ACTFL proficiency scales and testing techniques: Development, current status, and needed research. *Studies in Second Language Acquisition*, vol. 10. 129-47.
- Educational Testing Service. 1979. Manual for Peace Corps language testers. Princeton, N.J.: Educational Testing Service.
- Harvey, Philip R. 1968. Minimal passing scores on the graduate school foreign language tests. *Foreign Language Annals* 2.2:165-73.
- Lambert, Richard D., and Charles W. Stansfield. 1988. Development of a procedural model and associated testing instruments for the diagnostic measurement of language proficiency growth, attrition, and maintenance. Washington, D.C.: Center for Applied Linguistics.
- Omaggio, Alice C. 1980. Priorities for classroom testing for the 1980s. In: Dale L. Lange, ed., *Proceedings of the National Conference on Professional Priorities*. Hastings-on-Hudson, N.Y.: American Council on the Teaching of Foreign Languages. 47-53.
- Wilds, Claudia P. 1975. The oral interview test. In: Randall L. Jones, and Bernard Spolsky, eds., *Testing language proficiency*. Arlington, Va.: Center for Applied Linguistics. 29-44.

Testing English as a world language: Issues in assessing nonnative proficiency

Peter H. Lowenberg
Georgetown University

The past few years have been marked by considerable sophistication in the refinement of English language testing. Among their many advances, researchers and practitioners have successfully begun to incorporate a large number of learning, teaching, and contextual variables in the design and evaluation of test instruments. However, to date, relatively little attention has been focused on the variability that is inherent in the linguistic 'norms' for English that are generally tested. For in identifying these norms, most researchers in testing appear to assume implicitly that the benchmark for proficiency in English around the world should be the norms which are accepted and used by 'native speakers' of English.

This paper claims that such an assumption is no longer universally valid. In support of this claim, data are presented from the 'nonnative' varieties of English (Kachru 1986; Lowenberg 1986a), which have developed in countries formerly colonized by Britain or the United States where English is still used as a second, often official, language in a broad range of *intranational* domains, including government, commerce, the mass media and education. Such former colonies where English continues to function in one or more of these domains include Bangladesh, Botswana, Brunei, Burma, Cameroon, Ethiopia, Fiji, Gambia, Ghana, India, Israel, Kenya, Lesotho, Liberia, Malawi, Malaysia, Malta, Mauritius, Namibia, Nauru, Nigeria, Pakistan, the Philippines, Seychelles, Sierra Leone, Singapore, South Africa, Sri Lanka, Sudan, Swaziland, Tanzania, Tonga, Uganda, Western Samoa, Zambia, and Zimbabwe (Conrad and Fishman 1977; Encyclopedia Britannica 1986).¹

In these settings, English is used daily by nonnative speakers in the absence of native speakers, in non-Western sociocultural contexts, and in constant contact with other languages in multilingual speech communities. As a result, it often undergoes what Kachru (1986) has termed 'nativization'. That is, it develops new linguistic features at all linguistic levels, which would be considered deviant if used in countries where the 'native speaker' varieties of English (eg. British or American English) predominate. However, in their nonnative contexts, these linguistic innovations and modifications are so systematic, that many have become de facto local norms for English usage. In fact, attitudinal research (Kachru 1976; Shaw 1981) indicates that in at least two countries with widely attested nonnative varieties, India and Singapore, approximately 50% of college educated English users believe that nativized features should be the local norms for English usage and the models for English language teaching.

This paper proposes an approach for beginning to identify these norms in one nonnative variety: Malaysian English. A brief survey of the sociolinguistic context in which English has developed and is used in Malaysia is followed by an analysis of divergences from the norms of American English in sample texts written by Malaysians in several domains of 'Standard' English. Based on Trudgill (1983) and Tay and Gupta (1983), the standard model of a variety of English--native or nonnative--is here defined as the linguistic forms of that variety that are normally used in formal speaking and writing by speakers who have received the highest level of education available in that variety. Standard English is the accepted model for official, journalistic, and academic writing and for public speaking before an audience or on radio or television.

The focus in this analysis is on morphosyntactic features in Standard English because these features can be easily identified and classified for cross-varietal comparison, because they have already been well described in British and American English, because authoritative prescriptive norms for them are frequently available in such sources as school textbooks and newspaper style sheets, because proficiency in these forms is measured in most standardized tests of English proficiency, and because many of these features result from processes which are very productive in all varieties of English.² Analysis is of written, as opposed to spoken, texts since regional phonological rules, such as word-final consonant cluster simplification, can often mask the realization of morphosyntactic norms, and since written language has a greater likelihood of being successfully monitored or edited, making possible a distinction between sporadic 'mistakes' and systematic acquisitional 'errors' (see Corder 1967 [1974]).

Where relevant, also included are examples of Standard British English which diverge from American norms in ways similar to Malaysian English; of Standard Singapore English, used by Malaysia's southern neighbor; and of recent innovations in registers of Standard American English. Sources of data are major English language newspapers in England, the United States, and Malaysia; scholarly texts written by British, American, and Malaysian linguists; the style sheet for the Singapore *Straits Times*, the parent publication of Malaysia's preeminent English language newspaper, the *New Straits Times*; Malaysian government publications; and ESL textbooks authorized by the Malaysian Ministry of Education.

This analysis is followed by suggestions as to how familiarity with possible nativized features in Malaysian English can be used to distinguish 'varietal differences' from personal 'acquisitional deficiencies' in assessing the English proficiency of Malaysian speakers of English.

Malaysian English: Past and present. Malaysian English began to develop during the British colonization of the Malay Peninsula and western Borneo from the late eighteenth through the mid-twentieth centuries. The British established schools, especially in the urban trading centers on the west coast of the Malay Peninsula, where English was first taught and then used as the medium of instruction and of other school activities. The privileged recipients of this education--children of elite families from the indigenous Malays and from large South Asian and Chinese immigrant populations--came to use English increasingly in their daily affairs.³ As a result, when the British

began to withdraw from the area in the late 1950s, English had become the dominant language of the non-European elites, both as a language of power and prestige and as an interethnic link language (Le Page 1962; Platt and Weber 1980; Asmah 1985).

At the time of its independence in 1957, the Federation of Malaya retained English as one of two official languages, along with the major indigenous language, Malay, which was also selected as the sole national language. English continued to serve as the dominant language of education, government and law, and large-scale banking and commerce (Le Page 1962). The second half of the 1960s brought implementation of a policy to increase the use of Malay in the expanded nation of Malaysia. In 1969, the Ministry of Education initiated a policy whereby all English-medium schools were to become Malay-medium (Llamzon 1978; Platt and Weber 1980). By 1983, this process had been completed nationwide at the primary through tertiary levels of education (Lowenberg 1988).

One result of this language policy has been a dramatic increase in Malaysians' use of Malay, renamed *Bahasa Malaysia*, and a corresponding decrease in their use of English, particularly outside of the cities (Lowenberg 1986b). This diminishing use of English, especially in the schools, has produced widespread popular concern that the general level of English proficiency among Malaysians is in decline. The English language newspapers frequently run such headlines as 'First Aid Needed for Our English' (*New Sunday Times*, 4/21/85) and 'Decline and Fall of the English Language' (*New Straits Times*, 5/11/85).

However, such preoccupations with an overall decline in English proficiency among Malaysians overlook the fact that English is still widely used by the current Malaysian elites, who were educated in English-medium schools during the colonial or initial postcolonial periods and are still quite proficient in English.⁴ Augustin (1982) reports that of 1.1 million Malaysians who completed secondary school between 1956 and 1970, almost 70% had an English-medium education. As a result, Augustin concludes, 'the educated... Malaysians now in the prime of life and who play leadership roles in government and trade are fairly competent in the use of English...' These English-educated elites are concentrated in the urban centers of the Malay Peninsula, where approximately 25% of the population use English 'extensively... for intergroup communication' (Augustin 1982:251-52). They still set the standards for English usage as leading journalists for the most prestigious English language newspapers, and as writers and editors of the government authorized English language textbooks (Lowenberg in press). The following morphosyntactic examples of Malaysian English are taken from those sources.

Morphosyntactic innovations in Malaysian English. The following Malaysian examples reflect the same types of divergences from Standard American English that occur in Standard British English. Examples (1) through (3), from reputable sources of Standard British English, indicate a major process by which British and American English often differ: the treatment as count nouns of mass nouns which consist of countable units.

- (1) Some small initial fall-off in attendances is unavoidable. (*Times of London*, 10/27/86:17, in Algeo 1988:7)
- (2) ... iceberg lettuces are down in price and should be selling for between 35p and 55p, depending on size. (*Daily Telegraph*, 8/9/85:6, in Algeo 1988:7)
- (3) The teacher records a set of advertisements for entertainments. (Porter and Roberts 1987:185)

All of the underscored items would be incorrect in contemporary Standard American English. However, items (4) and (5) demonstrate that this process is also productive in American English, especially when it is extended by writers whose scholarship is highly regarded, such as the noted American linguists Charles Ferguson, Shirley Brice Heath, and Jean Berko Gleason.⁵

- (4) Equally certainly, twenty-five authors and two editors do not know enough to write this book, and by virtue of knowledges and viewpoints they may not provide as cohesive a book as a single author. (Ferguson and Heath 1981:xxviii)
- (5) Parents' eagerness to teach their 6-month-old children the prelinguistic routine 'bye bye' is one evidence of their desire to show that their baby is on its way to being a socialized person. (Berko Gleason 1988:276)

A Malaysian linguist of equivalent stature in Southeast Asia, Asmah Haji Omar, has likewise extended this process in (6).

- (6) In this context, there were variations such as (code) switchings between English and their own language. (Asmah 1985:20)

Further Malaysian extensions of this process in domains of Standard English occur in (7), from the front page of the leading daily English newspaper, and in (8), from an ESL textbook published by the Ministry of Education.

- (7) Complaints of threats and intimidations have surfaced and these could affect the security situation in the State... (*New Straits Times*, 5/1/86:1)
- (8) A consideration for others is most important. (Koh and Leong 1976:238)

Differences across varieties of Standard English also frequently develop in fixed collocations of verbs with particles and prepositions, as in (9) and (10), again from credible Standard British English sources. The underscored items here would be deviant in Standard American English.

- (9) This envisaged 16 to 20 'technology schools' in big cities, each catering for 1000 selected pupils... (*Times of London*, 9/15/86:1, in Algeo 1988:25)
- (10) The learner will only be able to show that his 'knowledge' of the text is approximating to that of the teacher through tests, reproduction, and answers to 'higher inference' questions. (Porter and Roberts 1987:182)

Similar constructions in Malaysian English occur in examples (11) through (14), from an Oxford University Press ESL text, a government publication, and two major English language newspapers.⁶

- (11) Give your book in. (Howe 1974:125)
- (12) ... Mr. Pierre MacDonald said at a press conference in Kuala Lumpur on April 30 to round up a four-day visit, that... (*Malaysian Digest*, May 1987:9, publication of Ministry of Foreign Affairs, Malaysia)
- (13) That way the forms would be filled and processed within minutes, rather than have the passengers fill up all the details while at the checkpoint. (*The Sunday Star*, 3/31/85:2)
- (14) With three days to go before acceptance, the battle for Umno Youth Exco seats is hotting up. (*New Straits Times*, 6/14/80:1)

A third type of cross-varietal morphosyntactic divergence occurs in prepositional collocations, as in (15), produced by the internationally respected British linguist John Swales.

- (15) Yet we appear to have a multi-targeted arrow to hand that will cover the wide ground of interest we are seeking... (Swales 1985:218)

Examples (16) and (17), from the style sheet of the Singapore *Straits Times*, illustrate how such differences from Standard American English occur in street locations in nonnative varieties of Standard English.

- (16) She lives in 6th Avenue. (*Straits Times Press*, 1985:4)
- (17) I live in an apartment at Belmont Road. (*Straits Times Press*, 1985:177)

These Singaporean norms may be following norms of British English (cf. Schur 1987; Algeo 1988). At any rate, the same system applies in Malaysian (18), where *jalan* is the Malay equivalent of English *street*.

- (18) A woman lost an envelope containing about \$13,000 to a snatch thief in Jalan Bandar Raya today. (*New Straits Times*, 5/7/86:4)

Differences or deficiencies? In this type of analysis, it is crucial to distinguish possible ‘innovations’, which result from such extensions of productive linguistic processes as those in the preceding examples, from features which are clearly ‘mistakes’ in any variety of Standard English, in that they do not result from these processes. An example of such a mistake in a domain of Standard American English occurs in (19).

- (19) But a U.S. district judge in New York temporary banned railroad machinists from honoring picket lines... (*Washington Post*, 3/6/89:A4)

It is likewise important to limit sources of items proposed as normative for a variety to those likely to have been written by the most highly proficient speakers of that variety. Not all English speakers who write in the domains of Standard English have attained this level of proficiency. For example, the underscored items in (20), taken from a Malaysian English language newspaper published on the island of Borneo, appear to reveal personal deficiencies in the writer’s and/or editor’s acquisition of English, rather than differences across varieties of Standard English.⁷

- (20) The sources feared that delegates in Saturday’s meeting may proposed for a 10-cent increases. (*Sabah Times*, 6/27/80:16)

Even in the prestigious *New Straits Times*, advertising copy, which is not always written or carefully proofread by highly proficient English speakers, may reflect such deficiencies, as in (21).

- (21) You can choose from the different types of floors’ layouts of our Triple Storey Shop Office... (Advertisement in *New Straits Times*, 5/7/86:13)

Although *floors’ layouts* may be analogous to British English *drugs overdose* or *departures lounge*, the use of the possessive here is not based on such a productive process.

In other cases, determination of acceptability can be more difficult. For example, the student writer on the Indiana State University student newspaper who uses *violated against* (22) three times in the same article may not be an unimpeachable reflection of Standard American English.

- (22) Morris claims his First Amendment rights were violated against when an ISU Safety and Security Officer demanded Morris give him his sign. (*The Indiana Statesman* [Indiana State University, Terre Haute], 2/16/89:1)

Malaysian instances of such questionable acceptability occur in (23) and (24), taken from the 'Letters to the Editor' section of the *New Straits Times*, and an advertisement reproduced in an Oxford University Press ESL textbook.

- (23) Never in my wildest imaginings did I think that you could come up with such a superbly managed show... ('People's Page,' *New Straits Times*, 6/22/83:15)
- (24) Take up diets you can trust... and exercises that are good to you. (Magazine advertisement reprinted in Vaz 1978:35)

All of the underscored items in (22) through (24) could result from the same productive processes that yield examples (1) through (14). However, in (22) through (24), we can surmise less about the writers' proficiencies in Standard American or Standard Malaysian English than we can in (1) through (14).⁸

Yet another problem in identifying Standard English in nonnative varieties is that ESL textbooks not published by a major press or the Ministry of Education may not accurately reflect English norms. Tan (1982), the source of example (25), is such a Malaysian English textbook.

- (25) One day I went to the roadside stall to buy some durians. I took up one and sniffed at it. (Tan 1982:56)

Ultimately, however, even if we can control for productivity of linguistic processes underlying divergences and for the English proficiency of sources, it is impossible to determine from data as limited as these whether an item is a token of a developing varietal norm or is simply a nonce innovation by an educated Malaysian English speaker. Clearly, a much broader data base, such as the Brown Corpus or the Survey of English Usage, is called for.

Implications for assessing English proficiency. Meanwhile, even such preliminary studies as this one can have important implications for testing nonnative speakers' proficiency in English. Most obviously, such analyses suggest limits on how far it can be assumed that norms of Standard American English extend to other varieties of Standard English, native or nonnative.

One practical implication of this observation concerns the international validity of certain items in the Test of English for International Communication (TOEIC), which the Educational Testing Service (ETS) has been administering since 1979. TOEIC is 'designed to test the English language as it is used internationally in business, commerce, and industry' (Educational Testing Service 1987:2). ETS points out that the TOEIC differs from the Test of English as a Foreign Language (TOEFL), which 'is designed to determine how well a candidate can use English in colleges and universities in the United States' (Educational Testing Service 1986:3). Thus, whereas the TOEFL is based on the norms of Standard American English, the TOEIC is implicitly unbiased toward any variety of Standard English.⁹

However, item (26), considered incorrect in a commercially published practice book for the TOEIC, would be acceptable in Standard Singapore or

Malaysian English (see (16) through (18) above), as well as in Standard British English (cf. Schur 1987:408; Algeo 1988:13).

- (26) He lives in Main Street. (Lougeed 1986:13)

Cross-varietal differences likewise become a factor in (27), a practice problem in the official ETS bulletin for the TOEIC. The student's task here is to identify the underlined item that is ungrammatical.

- (27) Please fill out the enclosed form to tell us how you think about our service. (Educational Testing Service 1987:18)

The ungrammatical item here is *how*. However, the construction *fill out* might well also be unacceptable to a student who had learned to *fill up* a form in Malaysia, as in (13) above, or to *fill in* or *fill up* a form in Standard British English (Trudgill and Hannah 1985:58; Schur 1987:135).

Examples (28) and (29), similar in examinee task to (27), are items that appeared in a TOEIC test (Educational Testing Service 1980:27-28).

- (28) His proposal met with a lot of resistances.

- (29) The new equipments shipped from Hong Kong will be the only items on sale this week.

Resistances and *equipments*, which result from the same productive process yielding (1) through (8) above, may well be acceptable to educated speakers of Malaysian or other nonnative varieties of Standard English.

Another implication of the development of nonnative norms concerns assessments of the English proficiency of the many foreign students enrolled in American universities who come from countries such as Malaysia, where nonnative varieties have been linguistically and attitudinally identified.¹⁰ Evaluators of these students' English proficiency can attempt to distinguish 'deficiencies' in a student's general English proficiency from 'differences' in the student's usage that diverge systematically from Standard American English as a result of highly productive morphosyntactic processes and might therefore be considered Standard English by highly educated English users in the student's home country.

Analysis of (30) and (31), taken from entrance examination essays written by Malaysian graduate students in linguistics at Georgetown University, illustrates how such a distinction might be made.

- (30) But most of all, it helped so much when I started to enroll into several writing courses in George Mason University.

- (31) I began to follow their advices. One of my instructor in George Mason University said that when a person learns a language he also learns the culture of the language.

All of the underscored items in (30) and (31) violate the norms of Standard American English and would most likely be considered incorrect by American university instructors. However, the constructions *enroll into* and *advices* could both result from productive processes illustrated in (1) through (14), and could thus possibly be acceptable in Standard Malaysian English. *In George Mason University*, which occurs in both (30) and (31), might likewise reflect correct prepositional usage in Standard Malaysian English, in a manner analogous to (18). In contrast, *one of my instructor* in (31) clearly does not result from any productive process in English, and is therefore most likely a mistake or an error in all varieties of Standard English. Such analyses of students' writing can help evaluators to distinguish between these students' overall deficiencies in Standard English and minor morphosyntactic adjustments which they must make in order to conform to the norms of Standard American English during their studies in the United States.

Conclusion. This paper has argued that the tradition of limiting the norms for Standard English to those of its native speakers is no longer universally valid. It has also proposed one approach for identifying morphosyntactic norms in Standard English that applies to both the native-speaker and the nonnative varieties of English.¹¹

This identification of norms in nonnative varieties of English is an area of linguistics in which research is still just beginning. Recently published studies have investigated nativization processes not only in morphology and syntax, but at other linguistic levels as well, including discourse (see, for example, Kachru 1985 and papers in Smith 1987). Accurate analysis and description of these norms presents formidable challenges for researchers, both conceptually and methodologically.

Meanwhile, however, insights from limited studies such as this one can have practical applications for improving the design of proficiency tests in English as a nonnative language and for analyzing nonnative speakers' deviations from native-speaker norms. On a more theoretical level, such research can be useful in writing a truly comprehensive grammar of English as a world language, and should yield valuable insights on basic processes of language variation and change.

Notes

- Richards (1979) has estimated that the total number of users of non-native varieties of English around the world may be almost as large as the total number of native speakers. While this claim is most likely exaggerated, the number of speakers of nonnative varieties is certainly quite substantial. For example, Kachru (1984) estimates that 27.4 million Indians regularly use English, giving India the third largest population of English speakers in the world after the United States and Britain.

- For discussions of nativization processes at other linguistic levels in nonnative varieties in general, and in Malaysian English in particular, see Platt, Weber, and Ho (1984), Kachru (1986), Lowenberg (1986a; 1986b; in press), and Smith (1987).

- The Malays, Chinese, and South Asians remain present-day Malaysia's

largest ethnic groups. In 1985, the ethnic distribution of Malaysia's total estimated population of 15,676,700 was 55.8% Malay and 33.3% Chinese; of the 10.9% 'others', the majority were South Asian (Encyclopedia Britannica 1986:736).

4. Le Page (1984:118) observes that 'the urban middle classes of all ethnic groups try to retain English'; even the urban middle-class Malays, who stand to benefit most from the shift to Bahasa Malaysia, 'may go to some lengths to ensure that their children are bilingual [in English] too, since a good command of English is still an advertised requirement for highly paid jobs, at least in the commercial sector.'

In contrast, as Abdullah Hassan (personal communication) points out, many young Malaysians with only limited proficiency in English come from rural areas, where until a recent expansion in the provision of universal education, little if any instruction in English was available. Therefore, the 'deterioration' in the Malaysian population's proficiency in English may not be as severe as some observers lament.

5. It is noteworthy that (3), as well as (10), produced by two British linguists, appear in a volume that was published, and therefore tacitly accepted, by an American publisher.

6. Particularly interesting is (14), *hotting up*, which I have now read on three occasions in the *New Straits Times*.

7. The constructions in *Saturday's meeting* and *proposed for* in (20) could reflect extensions of productive processes of the types manifested in (9) through (18). See discussion of (22) through (24).

8. In addition, the writers of (22) through (24) probably do not have sociolinguistic status as innovators equivalent to that of, for example, the authors of (4) through (6).

9. Queried by telephone, the TOEIC Program Office at Educational Testing Service stated that the norms used for TOEIC are from native-speaker varieties and do not include nonnative varieties.

10. According to data compiled by the Institute for International Education (reported in Magner 1989), in the school year 1987-88, 356,190 foreign students were enrolled in U.S. junior colleges, colleges, and universities in the United States. Of this total, more than 90,700--over 25%--came from the countries listed above as using English as a second language with some official status and from the British colony of Hong Kong, where English is an official auxiliary language. Moreover, among the fifteen polities with the most foreign students in the United States in 1987-88, five--India (21,010), Malaysia (19,480), Hong Kong (10,650), Nigeria (8,340), and Pakistan (6,570)--have widely attested nonnative varieties of English.

11. Also necessary for the identification of nonnative norms in a particular variety will be judgments of the acceptability of specific innovations in that variety by that variety's most highly educated speakers.

References

- Algeo, John. 1988. British and American grammatical differences. *International Journal of Lexicography* 1(1):1-31.
 Asmah Haji Omar. 1985. Patterns of language communication in Malaysia. *Southeast Asian Journal of Social Science* 13(1):19-28.

- Augustin, John. 1982. Regional standards of English in Peninsular Malaysia. In: John Pride, ed., *New Englishes*. Rowley, Mass.: Newbury House. 248-58.
- Berko Gleason, Jean. 1988. Language and socialization. In: Frank K. Kessel, ed., *The development of language and language researchers: Essays in honor of Roger Brown*. Hillsdale, N.J.: Lawrence Erlbaum Associates. 269-80.
- Conrad, Andrew, and Joshua Fishman. 1977. English as a world language: The evidence. In: Joshua Fishman, Robert Cooper, and Andrew Conrad, eds., *The spread of English*. Rowley, Mass.: Newbury House. 3-76.
- Corder, S. Pit. 1967. The significance of learners' errors. *International Review of Applied Linguistics* 5(4):161-70. Reprinted in Jack C. Richards, ed., 1974, *Error analysis: Perspectives on second language acquisition*. London: Longman. 19-27.
- Educational Testing Service. 1980. *Test of English for International Communication*. Princeton: Educational Testing Service.
- Educational Testing Service. 1986. *Guide for TOEIC users*. Princeton: Educational Testing Service.
- Educational Testing Service. 1987. *Bulletin of information, Test of English for International Communication*. Princeton: Educational Testing Service.
- Encyclopedia Britannica. 1986. *Britannica Book of the Year*, 1986. Chicago: Encyclopedia Britannica, Inc.
- Ferguson, Charles A. and Shirley Brice Heath. 1981. Introduction. In: Charles A. Ferguson and Shirley Brice Heath, eds., *Language in the USA*. Cambridge: Cambridge University Press. xxv-xxxviii.
- Howe, D.H. 1974. *New guided English, Book 1*. Kuala Lumpur: Oxford University Press. [Reprinted by Penerbit Fajar Bakti Sdn. Bhd., Kuala Lumpur, 1986]
- Hua Wu Yin. 1983. Class and communalism in Malaysia. London: Marram Books.
- Kachru, Braj B. 1976. Models of English for the Third World: White man's linguistic burden or language pragmatics? *TESOL Quarterly* 10(2):221-39.
- Kachru, Braj B. 1984. World Englishes and the teaching of English to non-native speakers: Contexts, attitudes, and concerns. *TESOL Newsletter* 18:25-26.
- Kachru, Braj B. 1986. *The alchemy of English*. Oxford: Pergamon Press.
- Kachru, Yamuna. 1985. Discourse analysis, non-native Englishes and second language acquisition research. *World Englishes* 4(2):223-32.
- Koh, Judy, and Leong Suat Chin. 1976. *Dewan's communicative English, Book 4*. Kuala Lumpur: Dewan Bahasa dan Pustaka, Ministry of Education.
- Le Page, Robert B. 1962. Multilingualism in Malaya. In: *Symposium on multilingualism [Proceedings of the second meeting of the Inter-African Committee on Linguistics, July 16-21, 1962, Brazzaville.]* London: Committee for Technical Cooperation in Africa.
- Le Page, Robert B. 1984. Retrospect and prognosis in Malaysia and Singapore. *International Journal of the Sociology of Language* 45:113-26.
- Llamzon, Teodoro A. 1978. English and the national languages in Malaysia, Singapore, and the Philippines: A sociolinguistic comparison. *Cross Currents* 5(1):87-104.
- Loughheed, Lin. 1986. English, the international language. Tokyo: Addison-Wesley.
- Lowenberg, Peter H. 1986a. Non-native varieties of English: Nativization, norms, and implications. *Studies in Second Language Acquisition* 8(1):1-18.
- Lowenberg, Peter H. 1986b. Sociocultural context and second language acquisition: Acculturation and creativity in Malaysian English. *World Englishes* 5(1):71-83.
- Lowenberg, Peter H. 1988. Malay in Indonesia, Malaysia, and Singapore: Three faces of a national language. In: Florian Coulmas, ed., *With forked tongues: What are national languages good for?* Ann Arbor, Mich.: Karoma Publishers. 146-178.
- Lowenberg, Peter H. (in press). The nativization of English in Malaysia: A case study in language pragmatics. In: Jenny Cheshire, ed., *English around the world: Sociolinguistic perspectives*. Cambridge: Cambridge University Press.
- Magner, Denise K. 1988. Students from Asia made up more than half of all foreigners at U.S. colleges last year. *The Chronicle of Higher Education*, October 26, A39.
- Platt, John, and Heidi Weber. 1980. English in Singapore and Malaysia. Kuala Lumpur: Oxford University Press.
- Platt, John, Heidi Weber, and Ho Mian Lian. 1984. *The new Englishes*. London: Routledge and Kegan Paul.
- Porter, Don, and Jon Roberts. 1987. Authentic listening activities. In: Michael H. Long and Jack C. Richards, eds., *Methodology in TESOL*. New York: Newbury House. 177-87.
- Richards, Jack C. 1979. Rhetorical and communicative styles in the new varieties of English. *Language Learning* 29:1-25.

- Schur, Norman W. 1987. British English, A to Zed. New York: Facts on File Publications.
- Shaw, Willard. 1981. Asian student attitudes toward English. In: Larry E. Smith, ed., English for cross-cultural communication. New York: St. Martin's Press. 108-22.
- Smith, Larry E., ed. 1987. Discourse across cultures: Strategies in World Englishes. New York: Prentice Hall.
- Straits Times Press. 1985. Our style: The way we use words in The Straits Times. Singapore: Straits Times Press, Ltd.
- Swales, John. 1985. ESP—the heart of the matter or the end of the affair? In Randolph Quirk and Henry Widdowson, eds., English in the world. Cambridge: Cambridge University Press. 212-23.
- Tan Chor Whye. 1982. Sistem kursus ulangkaji S.P.M., The new communication English, Subject 122. Kuala Lumpur: Pustaka Sistem Pelajaran Sdn. Bhd.
- Tay, Mary W. J., and Anthea Fraser Gupta. 1983. Towards a description of Standard Singapore English. In: Richard B. Noss, ed., Varieties of English in Southeast Asia. Singapore: SEAMEO Regional Language Centre. [Anthology Series 11] 173-189.
- Trudgill, Peter. 1983. Sociolinguistics (revised edition). Harmondsmith, England: Penguin Books.
- Trudgill, Peter, and Jean Hannah. 1985. International English: A guide to varieties of Standard English (second edition). London: Edward Arnold.
- Vaz, Fleur. 1978. English communication for syllabus 122. Kuala Lumpur: Oxford University Press.

Oral proficiency in the less commonly taught languages: What do we know about it?

Richard T. Thompson
Center for Applied Linguistics

In this paper I will attempt to summarize what is known about the application of foreign language proficiency testing activities to the less commonly taught languages (LCTLs). In so doing I will assume some knowledge of the proficiency movement on the part of the listener. I will only provide a fuller discussion of technical aspects if necessary to illustrate a point in the paper.

What do we know about proficiency testing in the less commonly taught languages?

- (1) We know something about its origins.
- (2) We know about its impact on the generic guidelines.
- (3) We know what some of the current problems are and some of their solutions.
- (4) We know a little bit about future demand.

1 The origins. Although Titone (1968) traces the quest for proficiency back to 3000 B.C. and the first historical description of a proficiency test has been described in a biblical story involving the pronunciation of the word *shibboleth*, I will pick up the subject in this century and in this country.

The National Mobilization and Manpower Act of 1952 required the Civil Service Commission to develop a register of government employees with foreign language backgrounds. By 1956 the Foreign Service Institute (FSI) had prepared preliminary skill level descriptions and conducted a survey of over 4000 foreign service officers. The findings of the survey led to the announcement of a language policy which stated that foreign language skills were vital to the conduct of foreign affairs and led to a decision to verify whether a 'useful' skill level had been achieved (Liskin-Gasperro 1985).

FSI adopted a face-to-face interview procedure which, although based upon the original descriptions, added both linguistic and functional components. In the decade that followed the rating scale was adopted or endorsed by the member agencies of the Interagency Language Roundtable (ILR).

In the late 1960s the Peace Corps, which was teaching languages to thousands of trainees in the United States and abroad, had become aware of the work of the ILR and FSI in the area of proficiency testing. At the request of the Peace Corps, FSI agreed to test both trainees and volunteers at the beginning, at midpoint, and at the end of their training. FSI quickly realized

that it did not have the resources to continue such a large external testing program. The Educational Testing Service (ETS) agreed to assume these testing responsibilities. After training by FSI, ETS undertook the first large-scale interview testing activity outside the U.S. government.

During 1977-78, with support from the National Endowment for the Humanities and the Rockefeller Foundation, the Modern Language Association (MLA) and the American Council of Learned Societies (ACLS) convened the *MLA/ACLS Language Task Forces Project*. The task forces included: Institutional Language Policy, the Commonly Taught Languages, the Less Commonly Taught Languages, and Public Awareness and Government Relations.

The reports (Brod 1980) of the task forces were concluded in the fall of 1978 and were subsequently considered by the President's Commission on Foreign Language and International Studies. Among the recommendations of the reports relating to proficiency assessment were the following:

- Institutions should be encouraged to adopt nationally recognized performance or proficiency standards.
- The *MLA* and *ACLS* should secure funding for the revision and redevelopment of tests for the measurement of proficiencies in the four language skills in all the most commonly taught and wide-use languages.
- The *MLA* and *ACLS*, in association with other professional groups and testing organizations, should develop an outline of realistic proficiency goals by stages of achievement.

According to the reports, the purpose of the latter recommendation was to develop a set of universal standards similar to those used at FSI.

Also during this important period the FSI in 1979, with assistance from the U.S. Department of Education, launched the FSI Testing Kit Workshops. These workshops were to have a profound effect on the introduction of proficiency standards into the academic community. College professors were invited to participate in workshops which trained them in the application of the FSI oral proficiency interview techniques. One conclusion reached by the participants was that there was a need to expand the lower end of the scale to make the finer differentiations required in an academic setting which is typified by fewer hours of instruction stretched over longer periods of time, a setting in which student achievement rates barely begin to show up on the traditional scale.

In yet another separate but related activity, the ETS Common Yardstick Project funded by the U.S. Department of Education also concluded that the FSI/ILR rating scale, subject to modifications similar to those recommended by the Testing Kit Workshops, was suitable for use in an academic setting.

Senator Paul Simon (1978), taking seriously his role as member of the Commission on Security and Cooperation in Europe, urged President Carter to establish a commission to assess the response of the United States to the requirements of the Helsinki Agreement that the 34 signatory nations will encourage the study of foreign languages and civilizations.

In the spring of 1978, the President's Commission on Foreign Language and International Studies was formed. Its four major tasks were (see Burn 1980):

- to recommend how public attention should be directed to the importance of foreign language and international studies,
- to assess the need in the United States for foreign language and area specialists and the job market for such specialists,
- to recommend what foreign language and international education programs are appropriate at different levels, and
- to specify legislative changes needed to implement the commission's recommendations.

The report (Strength through Wisdom 1979), which was submitted to President Carter in the fall of 1979, called for the establishment of a National Criteria and Assessment Program to develop foreign language proficiency tests and report on, monitor, and assess foreign language teaching in the United States. The report specified that the program would establish language proficiency goals for each year of study, with special attention to speaking proficiency.

According to Liskin-Gasparro (1985:31):

The MLA-ACLS Language Task Forces' Reports and the President's Commission Report, with their recommendations that the profession establish proficiency-based course goals and construct proficiency tests to measure these outcomes, combined with the recommendation of the ETS Common Yardstick Project on the expansion of the lower end of the government oral proficiency scale and the enthusiastic reaction of the participants at the FSI Testing Kit Workshops set the stage for the appearance of the ACTFL Provisional Proficiency Guidelines in 1982.

The American Council on the Teaching of Foreign Languages (ACTFL) began work on the ACTFL/ETS guidelines in 1981 with a grant from the U.S. Department of Education. The initial project involved the development of guidelines for French, German, and Spanish, as well as the training of individuals to conduct testing in French, German, Italian, and Spanish.

It wasn't until 1984 that the ACTFL activities extended to the LCTLs with the addition of proficiency guideline development for Chinese, Japanese, and Russian and oral proficiency tester training in Arabic, Chinese, ESL/EFL, Japanese, Portuguese, and Russian. Subsequently, ACTFL, together with the Center for Applied Linguistics (CAL), designed and implemented a dissemination project to extend proficiency concepts to Hindi, Indonesian, and Swahili, as well as to provide some preliminary tester training in Hindi, Indonesian, Swahili, Hausa, and Lingala (Thompson, Thompson, and Hiple 1988).

2 LCTL impact on the generic guidelines. As the initial working committees began their work in Chinese, Japanese, and Russian, some discordant notes began to sound. It was becoming clear that the so-called 'generic' guidelines frequently reflected the Eurocentric bias of the earlier working committees in French, German, and Spanish. The problems were most evident in the statements concerning accuracy in speaking and with the writing systems (Hiple 1987).

The LCTL committees felt that learners of their languages had to deal with very different problems. References in the generic guidelines to inflections, articles, prepositions, subject-verb and adjective-noun agreement, which were made with reference to accuracy in speaking, did not fit many of the LCTLs.

The decision was made to remove many of the qualifying examples from the generic guidelines and place them in the specific language descriptions where they belonged. For example, in the Intermediate-Mid speaking descriptions, the references to subject-verb agreement, adjective-noun agreement, and inflections were removed in the revised guidelines of 1986. This left the LCTLs free to include accuracy statements which better represented their languages. For example, the Chinese guidelines refer to word order, auxiliaries, and time markers; the Arabic to verb-object phrases; the Japanese to formal nonpast/past and classifiers; and the Russian to a developmental hierarchy of cases.

The development of guidelines for languages with predominantly nonphonologically based writing systems such as Chinese and Japanese caused considerable debate and resulted in a number of important changes in the provisional guidelines. For example, distinction is made in the Chinese and Japanese reading guidelines between reading specially prepared materials and authentic texts and the reading of materials with or without the use of a dictionary.

In the discussion so far the first time I have made specific reference to the LCTL-specific guidelines has been in connection with the ACTFL guidelines, and especially their impact on the generic guidelines. This could leave one with the impression that prior testing using the scale either did not include the LCTLs or that it was done poorly. Nothing could be further from the truth. The early applications of the ILR/FSI oral proficiency interview (OPI) in the Peace Corps, along with its use of more than 30 years within Federal agencies such as FSI, DLI, and CIA, included extensive testing with the LCTLs.

Prior to the adaptation by ACTFL/ETS of the ILR/FSI guidelines, OPI testing was carried out extensively and successfully by the user agencies. It is necessary to understand that the existence of the generic guidelines and training in OPI are quite separate issues. The guidelines, whether ACTFL generic or ILR generic (there is only one guideline), lay a theoretical base for and provide a verbal description of the underpinnings of the OPI training. The guidelines as used by the ILR member agencies are nonlanguage-specific and have not impeded testing in the commonly taught or the LCTL languages. Furthermore, as the government agencies added new languages, not only were there no guidelines for these languages but there were no personnel trained in the testing of those languages to train new staff. Yet testing in these languages was carried out.

This experience both raises important questions about the role of the language specific guidelines and suggests valuable alternative strategies for OPI training today as it is carried out in the academic community.

3 Current problems. Problems in adapting the ACTFL generic guidelines to specific less commonly taught languages are of two kinds--theoretical and practical (Thompson, Thompson, and Hiple 1987).

3.1 Theoretical problems. Theoretical problems include the presence of code switching at high levels of proficiency among educated native speakers of Hindi, the early appearance of significant problems in register in Bahasa Indonesia and Japanese, the discordance of writing systems such as Chinese and Japanese with the reading and writing guidelines, and special problems of diglossia in Arabic. I have described some of these above. There are, of course, other problems of a theoretical nature, for example, in the research area, that have been raised, but these are not specific to the LCTLs and will therefore not be addressed here.

3.2 Practical and policy-related issues. Practical and policy-related issues include problems of tester training in languages where no trainers currently exist and the appearance of recent Federal legislation mandating competency-based language training and proficiency testing.

3.2.1 Practical issues. Tester training only became available in the LCTLs to the academic sector in 1984. Initial workshops were offered jointly by trained testers of the ILR and ACTFL testers and included Arabic, Chinese, Japanese, Portuguese, and Russian. The decision as to which language to include was made at least in part by the availability of a government tester to lead the training in that particular language. Early reliance on ILR testers has abated with the gradual training and certification of academic testers and trainers in all of the commonly taught languages and some of the LCTLs.

New federal legislation mandating proficiency testing has immediate and large-scale implications both for tester training and the development of language-specific guidelines for the nearly 160 LCTLs currently available at U.S. institutions of higher education with large foreign language and area studies programs funded by the U.S. Department of Education. Increasing demands for training are placing a serious strain on our national training capacity. For most of these languages no trainer is available, even within the federal government. It is likely that startup training for most of these languages will be through another language known to the prospective tester. Such training has already been initiated in English, French, German, and Russian for future testers of Hindi, Bahasa Indonesia, Polish, Hebrew, Hausa, and Thai.

An alternative testing strategy extensively used by U.S. government testers holds promise for use with the LCTLs. This procedure involves the presence of a certified tester (preferably in a related language) and an educated native speaker/informant of that language. The government experience in this area suggests that with appropriate direction and close observation of the native speaker during the administration of the test, reliable ratings can be assigned by such a rating team.

The development of semidirect tests of oral proficiency also holds promise in languages for which maintaining a cadre of trained testers is simply not feasible. Currently, the Center for Applied Linguistics has such tests available for Chinese and Portuguese and is planning on extending such development to Hebrew, Bahasa Indonesia, and Hausa.

3.2.2 Policy issues. Proficiency testing in the LCTLs presents challenges at the heart of several policy issues, some of which also face the commonly taught languages--issues involving the federal, state, and local governments, institutions of postsecondary education, the professional associations, and consumers in business and industry.

The President's Commission on Foreign Language and International Studies, previously mentioned, was charged with recommending legislation in the area of foreign languages. As it turned out, the timing of the Commission was compatible with the reauthorization cycle for the Title VI programs and the Commission was able to make a preliminary draft of its recommendations available to the congressional authorizing committee in time to have a profound affect on the Education Amendments of 1986.

These amendments included a number of highly significant changes in Title VI of the Higher Education Act (formerly Title VI of NDEA). This legislation is unique in many respects in mandating competency-based language training and testing for the nearly 150 programs receiving assistance under this Act. The U.S. Department of Education's Center for International Education, which administers these programs, has issued federal regulations which define competency-based language training to include validation through a testing procedure based on national standards.

These regulations create a new set of problems and a new urgency. As the affected universities seek to come into compliance with the new legislation, there will be intense competition for the limited training resources available. The Department of Education, institutions of higher education, and the major concerned professional associations will need to cooperate in setting realistic goals.

As a first step, ACTFL and the Center for Applied Linguistics, with partial assistance from the U.S. Department of Education, are cooperating in a survey of the foreign language and area studies centers and fellowships programs affected by the legislative changes. The survey is seeking information on the perceived need of these programs for guidelines development and training in proficiency testing as well as other areas. Which languages should be designated for priority development and who will receive initial tester training? Will interim procedures need to be developed to comply with federal regulations until a sufficient cadre of trained testing specialists becomes available?

Other policy questions that derive from the legislation ask what the role of the new federally authorized language resource centers will be. Could a small number of regionally located centers assume responsibility for testing and tester training in their regions? How would the relationships and responsibilities of the National Resource Centers and the privately funded National Foreign Language Center at Johns Hopkins be defined and coordinated? Will training be extended for the LCTLs to the precollege level and will teachers at that level receive training?

4 Future demand. What do we know about future demand? Results from the survey I mentioned above are now becoming available. Although it is still too early to present statistical results, it is becoming clear that there is widespread recognition among the institutions affected by the new federal

regulations that these requirements are mandated by law and must be complied with. Furthermore, many respondents acknowledge that important self-evaluation of language programs is taking place as a result of the requirements and that this is beneficial.

References

- Brod, Richard I., ed. 1980. Language study for the 1980s: Reports of the MLA-ACLS task forces. New York: The Modern Language Association.
- Burn, Barbara B. 1980. The President's Commission on Foreign Language and International Studies: Its origin and work. *Modern Language Journal* 64:7-8.
- Hipple, D. 1987. The extension of language proficiency guidelines and oral proficiency testing to the less commonly taught languages. Paper presented at the Indiana Symposium on Foreign Language Evaluation, Bloomington.
- Liskin-Gasparro, Judith E. 1985. The ACTFL proficiency guidelines: A historical perspective. In: *Teaching for proficiency, the organizing principle*, ed. T.V. Higgs. Lincolnwood, Ill.: National Textbook Co.
- Simon, Paul. 1978. Language and national policy. In: *Language in American life*, ed. E. Michael Gerli, James E. Alatis, and Richard I. Brod. Washington, D.C.: Georgetown University Press.
- Strength through Wisdom: A critique of U.S. capability. A report to the President from the President's Commission on Foreign Languages and International Studies. 1979. Washington, D.C.: U.S. Government Printing Office.
- Thompson I., R. T. Thompson, and D. Hipple. 1988. Issues concerning the less commonly taught languages. In: *Second language proficiency assessment: Current issues*, ed. Pardee Lowe, Jr. and Charles W. Stansfield. Englewood Cliffs, N.J.: Prentice-Hall.
- Titone, Renzo. 1968. *Teaching foreign languages: An historical sketch*. Washington, D.C.: Georgetown University Press.

Language proficiency testing with limited English-proficient students

J. Michael O'Malley

Georgetown University

Evaluation Assistance Center (EAC) East, Director

Abstract. There have been dramatic shifts in views on language proficiency theory and testing over the past decade which have dated most of the procedures used for language assessment in instructional programs for limited English-proficient students. Because the impact of these shifts has not been felt in assessment practices, current procedures for testing language minority students may be identifying and diagnosing students inaccurately. This paper analyzes some of those shifts and identifies specific problems with current language proficiency assessment instruments. In recognition of these problems, some local school districts and researchers have developed alternative assessment procedures that are preferred to the conventional tests. The paper concludes by discussing some of these alternative procedures and points to future directions for language proficiency assessment.

There have been dramatic shifts in views on language proficiency theory and testing over the past decade that have dated most of the procedures used for language assessment in instructional programs for limited English-proficient students. These shifts have occurred in parallel with changes in curriculum and instruction for limited English-proficient students that have also influenced the types of tests that are needed and the way in which testing is used. These shifting views have produced a complicated pattern of demands on proficiency assessment that existing instruments have been unable to address.

This paper sketches the context for some of these changes, describes the nature of their influence on language assessment, and indicates some of the adaptations to these changing circumstances that have emerged in the field. The paper begins by reviewing the purposes of testing in instructional programs for limited English-proficient students. A number of the shifts in views of language proficiency theory and testing will then be delineated along with some of the changes in curriculum and instruction that have had an impact on proficiency assessment. Some of the resultant problems with language proficiency assessment will be described, and a number of alternative assessment approaches that have emerged in response to these problems will be presented. The paper concludes with a statement of needs for future research and development with assessment procedures for limited English-proficient students.

Purposes of testing with limited English-proficient students. The practice of language proficiency assessment with limited English-proficient students is in part a response to the needs of educational programs and in part a response to governmental funding requirements and legal decisions. Because of these diverse influences, the purposes of language assessment have varied considerably. Nevertheless, there are at least four purposes and perhaps a fifth that tend to be found in most instructional programs (Henning 1987):

- (1) Selection. Identification of language minority students who are entitled under court decisions or under state or federal laws to receive special forms of instruction based on their limited proficiency in English.
- (2) Diagnosis. Delineation of the pattern of strengths and limitations in language skills that can be addressed through instruction.
- (3) Placement. Identification of the appropriate level of instruction in English and type of program that will most benefit the student.
- (4) Reclassification. Identification of when the student may shift to a level of instruction in English that is more demanding or shift to an instructional program that is not adapted to individual language needs. and
- (5) Program evaluation: Identification of the amount of growth in language proficiency that is associated with participation in the instructional program.

The first four of these purposes are decisions made for individual students, whereas the last purpose (evaluation) provides information about groups of students. The last purpose is usually exercised only when the program is operating under external funding from either the state or the federal government, but may be used if there is a local interest in data on program effectiveness.

Instructional programs sometimes use different language proficiency instruments for different purposes. For example, the specialized nature of diagnostic assessment calls for a test that will yield information on individual language subskills in addition to the total score. However, a single language proficiency instrument is sometimes used for reclassification as well as initial selection. In these cases, the schools often use multiple criteria such as grades, ratings, and test scores in reading and math to determine the student's readiness to advance to a different level of instruction. The use of instruments for program evaluation often calls for tests that are normed on a large population or that have standard scores, which may not be necessary to serve the other purposes.

Shifts influencing test selection and use. Our views of how language proficiency is defined and how test information can be used have shifted dramatically in the past ten years. Some of these shifts have occurred

specifically in language proficiency assessment and some have occurred in testing in general.

We began the decade in 1980 with at least two major shifts in how language proficiency was defined largely resolved and a third shift emerging. First, there was a shift from viewing language proficiency as the sum of a number of discrete elements to viewing language as an integrative whole (e.g. Oller 1979). Second, language proficiency came to be viewed in terms of communicative skills, consisting of broadly defined grammatical, strategic, discourse, and sociolinguistic competencies (e.g. Canale and Swain 1980). The shift which emerged about the change of the decade held that a special type of language proficiency is needed by minority language students to succeed in school. This special type of language consists of decontextualized and cognitively demanding language, unlike the highly contextualized and nondemanding language used in most instructional programs for limited English-proficient students (Cummins 1982).

More recently, questions have been raised about the adequacy of language proficiency as a construct to predict the success (or explain the failure) of minority language students. It has been suggested that students need to develop schemata (or integrated memory structures) associated with specific content areas in order to succeed in school (Saville-Troike 1989). All of these changes have implications for how language proficiency is assessed. For example, the view that schemata are important places emphasis on the need to determine the level and type of content area knowledge students possess as well as their language proficiency.

In addition to these changing views of language proficiency, there have been other changes that influence the way in which assessment is used with limited English-proficient students. One is the change in curricular emphasis in English as a second language programs from an exclusive focus on language to a broader focus on teaching language through the content areas (Chamot and O'Malley 1987, Crandall and Willetts 1986). This particular shift calls for special kinds of instruments for limited English-proficient students that will assess the unique type of language used in different content areas as well as the knowledge conveyed in content areas. A second change is the shift from audiolingual methodology to more eclectic instructional approaches including language experience, whole language, and cooperative learning (e.g. Freeman and Freeman 1988, Jacob and Mattson 1989, Richard-Amato 1988). There should be an attendant shift in language proficiency assessment from identifying isolated language skills to gaining a broader understanding of a student's ability to convey meaningful utterances through speech and writing. A third change is our understanding of the learning process, which has shifted from focusing on the conditions that influence the learner to identifying the mental processes used by the learner (e.g. Gagne 1986; Jones, Palincsar, Ogle, and Carr 1987; O'Malley and Chamot 1989). The accompanying shift in assessment results in an effort to identify the extent to which learners have mastered the mental processing that is linked to learning as well as the extent to which they have mastered specific skills and information that are typically stated as the goals of the curriculum.

Another element in this complex array of influences on language proficiency assessment is the changing nature of policy demands in which

tests and curricula play a significant role. One example is the demand for testing associated with accountability in categorical programs funded at the state or federal level. Each categorical program may demand a different type of assessment, often leading to overtesting and a call for simplification. A second example is linked to national declines in test scores among all students. As a response to this decline, there is increased emphasis on testing for accountability to identify state and district profiles (Ellwein, Glass, and Smith 1988; Wise 1988), and there is increased attention given to establishing uniform state- or district-level curriculum requirements (e.g. Darling-Hammond and Wise 1985, Hirsch 1989). A third example is the increasing use of minimum competency testing in states and particularly the requirement to pass tests for high school graduation (Education Commission for the States 1985). Few of these testing requirements have been adapted to the needs of limited English-proficient students. What all of these influences add up to is a dramatically altered role for testing in education, and a demand to use tests more frequently and for a greater variety of purposes.

One example of the influence deriving from use of testing for accountability is the pressure placed upon superintendents, principals, and teachers to produce increased scores on nationally normed tests. As a result of this pressure, school personnel have become very interested in how to select tests that match their local curricula, how to use test results for program improvement, and how to use assessment procedures to monitor student progress throughout the school year (O'Malley 1988). This applies to teachers of limited English-proficient students as well as to teachers in the mainstream.

Problems with language proficiency instruments. With few exceptions, the more widely used tests that assess language proficiency were developed more than ten years ago. Because so many of the language proficiency measures were developed before this decade began, they provide information that is not consistent with current views of language proficiency, proficiency assessment, or with needs of educators in the field.

Two of the more widely used exceptions that were developed in this decade are the Language Assessment Battery (Abbott 1985) and the Maculaitis Assessment Program (Maculaitis 1982). Each instrument assesses all four language skills and is based on more current (but quite different) theories of language proficiency. The fact that they were developed more recently, however, does not mean that they meet the needs of all projects. A number of projects are quite aware of these instruments but, as will be seen, prefer to use other approaches.

The language proficiency measures used by many projects have three characteristics: (1) they focus on oral/aural language skills, (2) they have a single form, and (3) they provide scores without normative information. The scores reflect the student's standing on a 5-point or similar continuum from no proficiency to native-like proficiency. The problems with these instruments are as follows.

(1) Concentrating on oral/aural proficiency alone may be inconsistent with the needs of projects to have information on all four language skills and provides no information in the content areas, as many projects require for reclassification decisions.

(2) Having a single form of the instrument fails to reflect the need for repeated pre- postassessment of students to determine growth, as is necessary for program evaluation or for monitoring student progress.

(3) Using scores on a 5-point or similarly delimited scale does not reflect the full underlying continuum of language proficiency.

(4) Using scores without normative information leaves projects with no basis for making comparisons relative to other populations or to determine growth.

Many projects have recognized the difficulties with these instruments and have developed their own tests or have supplemented formal assessment with rating scales. Some of these efforts are described in the following section.

Alternative assessment approaches. Three specific instruments developed by local school districts for purposes of placement were selected for discussion in this paper. The first is the Computerized Adaptive Test of Minimum English Competency (CAT/MEC) (Gross and Stevenson 1986), which assesses reading and writing skills in English; the second is the Criterion-Referenced Test of Content Areas (Martinez and Cunningham 1989), which assesses knowledge and reading skills in content areas; and the third is the Secondary Integrated Language Tasks (Nelson 1988), which assesses all four language skills using an integrative language approach. The instruments were selected to represent a cross-section of the varied responses by local school districts to the need for new instruments. These instruments address only some of the general difficulties in language assessment identified earlier but seem to satisfy the needs of the local practitioners by whom they were developed.

The characteristics of each instrument are portrayed in Table 1, including the language skills assessed, the purpose of the instrument, the grade levels, item types, types of scores, levels of proficiency assessed, and other features of the instruments. What this table reveals is the diversity of the characteristics of the instruments and the fact that, for the same purpose (placement/reclassification), such varied types of instruments had been developed to meet local needs. Developers of each of these instruments, when asked, indicated that their test was designed to match the local curriculum, illustrating further the variation in local conditions and needs.

The computer adaptive test addresses concerns about overtesting because the actual testing time in CAT/MEC is reduced to a minimum. The criterion-referenced test of content areas addresses concerns about being able to identify native language content area knowledge at the time of placement, and being able to assess content area knowledge in English at the time of reclassification. The integrative language test has theoretical currency because language is assessed as an integrative whole while at the same time providing information for teachers to use in diagnosing each student's strengths and weaknesses in language. This instrument also addressed local concerns about eliciting information for decision making in as short a time as possible. What these instruments do not provide is national norms or even norms that go beyond the immediate school district where the test was developed. Furthermore, the test content may be valid only for the district curriculum where the instrument was designed. Thus, any district adopting the instrument used by another district would have to collect additional information.

Table 1. Characteristics of three placement tests.

Characteristic	CAT/MEC test	CRT Content test	Secondary ILT test
Language skills:	Reading, writing structures	Reading and reading in content areas	Listening, speaking, reading, writing
Purposes:	Placement/reclassification	Placement/reclassification	Placement/reclassification, diagnosis
Grade levels:	Piloted at secondary, but MEC is K-12	1, 2-3, 4-6, 7-12	Secondary
Item type:	Multiple-choice (MC), 4 alternative	MC, 4 alternative and cloze	Varies by test, but includes oral response, listening, MC, dictation, cloze
Scores:	One score	One score on the English test; Separate scores in math, social studies, and science in Spanish	Listening, 2-grammar, 2-reading, 2-writing (7 scores and 7 separate tests)
Level of proficiency:	Requires basic reading skills	Requires basic reading skills	All levels
Test materials:	CAT	Paper and pencil	Photographs, paper and pencil
Administration:	Individual	Group	Individual, group
Forms:	Adaptive	Three levels at each grade	One
Languages:	English	English, Spanish	English
Norms:	Piloted locally	Piloted locally	Piloted locally
Administration:	11 min. reading, 7 min. writing on the average	Untimed	10-20 min. writing, 12-15 min. language with grammar, 15 min. reading

In addition to these varied types of instruments that have been developed by local school districts, there has been increasing interest in the use of rating scales and checklists to monitor student progress. Some of these instruments have appeared in the research literature and others have evolved out of practitioner applications. A number of them have been developed or adapted by the Evaluation Assistance Center (EAC) East at Georgetown University, a project funded by the U.S. Department of Education to provide technical assistance to state and local education agencies on assessment and evaluation in instructional programs for limited English-proficient students. These instruments have at least three characteristics in common: a focus on measuring student capabilities rather than student deficiencies, the opportunity to obtain immediate feedback on student performance for use in instructional planning, and the potential for repeated use to assess student progress.

One example is shown in Table 2, the Literacy Development Checklist. The Literacy Development Checklist contains four sections: reading skills, interest level, applications, and reading strategies. The interest and applications sections reflect the whole language approach, the reading

strategies section reflects recent theoretical views on mental processes underlying reading, and the reading skills section provides information on the basic skills required to perform reading comprehension tasks.

Table 2. Literacy development checklist.

Student: School: Mark: x = Effective / = Sometimes effective - = Needs work	Teacher: Academic year:	Quarter: 1	2	3	4
Reading Processes:					
1. Reading skills: Comprehends oral stories Reading vocabulary Fluent decoding Literal comprehension in reading Inferential comprehension					
2. Interest Initiates own reading Shows pleasure in reading Selects books independently Samples a variety of materials					
3. Applications: Participates in language experience story development Participates in reading discussion groups Writes appropriate dialogue journal entries Chooses books of appropriate difficulty Uses reading in written communication					
4. Reading strategies: Monitors attention Notices miscues that interfere with meaning Infers meaning based on: Word clues Sentence structure Story structure Prior experience Summarizes main ideas or key events Links details to main ideas Remembers sequence of events Predicts conclusions Requests help if needed					

Note. Adapted for Georgetown University Evaluation Assistance Center (EAC) East from materials developed by the National Council of Teachers of English and by the Writing Lab of the University of New Hampshire.

A second example is shown in Table 3, the Interpersonal and Academic Language Skills Checklist. This checklist reveals some of the thinking in the 1980s about cognitive academic language skills that are required for limited English-proficient students to succeed in school (e.g. Cummins 1982). The four categories in which language proficiency is rated are based on the notion

that language proficiency can be defined along two continua: a dimension concerned with the cognitive demand for performing the task, and a dimension concerned with the degree to which language meaning is contextualized or supported by paralinguistic cues. Since these dimensions are bipolar, they define four cells as suggested by the four categories in Table 3, with the examples of language skills that are defined for each cell.

Table 3. Interpersonal and academic language skills checklist.

Name:	Date:
<p>Directions. Please check skills which have been observed at an appropriate level in either English or the non-English language.</p>	
Contextualized/Noncognitively demanding:	English Non-English language
1 Answers basic questions appropriately. 2 Exchanges common greetings. 3 Follows general classroom directions. 4 Participates in routine school activities. 5 Describes classroom objects or people. 6 Gives classroom commands to peers. 7 Participates in sharing time. 8 Retells a familiar story. 9 Initiates and maintains a conversation. 10 Follows along during oral reading.	
Decontextualized/Noncognitively demanding:	
11 Decodes fluently. 12 Reads noncognitively demanding information, e.g. notes, signs, directions, simple sentences, etc. 13 Writes words and simple sentences. 14 Generates simple sentences. 15 Writes from dictation.	
Contextualized/Cognitively demanding:	
16 Follows specific directions for academic tasks. 17 Uses terms for temporal and spatial concepts, e.g. <i>first, last; top, bottom; left, right</i> , etc. 18 Asks/answers questions regarding academic topics. 19 Understands contextualized academic content. 20 Reads stories for literal comprehension.	
Decontextualized/Cognitively demanding:	
21 Distinguishes main ideas from details (oral). 22 Predicts conclusions after listening to story. 23 Understands lectures on academic content. 24 Uses language to reason, analyze, synthesize. 25 Participates in academic discussions. 26 Reads content area information for comprehension. 27 Uses glossary, index, appendices, etc. 28 Writes meaningful short paragraphs. 29 Uses correct language mechanics. 30 Writes coherent stories or reports.	

Note. Adapted for Georgetown University Evaluation Assistance Center (EAC) East from 'Checklist for language skills' (Bernhard and Loera) and based on Cummins (1982).

The other checklists developed by the EAC East include a Writing Development Checklist, Thinking Skills Checklist, Language Functions Observation Form, Student Self-Rating Progress Checklist, Self-Rating Strategies Checklist, and Assessing What Students Know, some guidelines for teachers to elicit information about the knowledge students possess about any topic. Together, these checklists and rating forms represent a response to the diverse needs of school districts for alternative forms of assessment. What remains unresolved, and would need to be addressed by individual districts, is the way in which the instruments would be used, and the criteria that would be reached for some of the more ambiguous terms in the checklists such as 'comprehends,' 'fluent,' and 'appropriate.' These issues have been the subject of workshops presented by the Georgetown University EAC East.

Summary and conclusions. What does this all mean for assessment in instructional programs for limited English-proficient students? There are at least three major points that emerge from this discussion. The first is that existing instruments in the field have not satisfied the needs of all users, resulting in a demand for additional instrumentation that some projects have resolved locally. The second point is that local demands are varied and are influenced by a complex array of conditions that will be represented locally in different combinations. A third point is that the complex requirements for language proficiency tests are probably best addressed by instruments that focus on a limited number of school district and programmatic needs. However, because school district needs are so varied, and because there is a need to reduce overtesting, rating scales or checklists may be required that will serve multiple purposes, that do not require time drawn away from instruction, and that can be integrated with instructional planning.

There exist a variety of instruments and rating scales that are available for language proficiency assessment, monitoring student progress, and program evaluation. Projects should rely on multiple sources of information obtained through varied types of data collection, and thereby increase the accuracy with which they perform selection, placement, diagnosis, and evaluative functions. There is no large-scale programmatic research being performed currently that addresses language proficiency from a theoretical or an applied view. Thus, advances in measurement procedures are likely to emerge from small projects such as those identified, where local practitioners are struggling to meet programmatic needs.

References

- Abbott, M.M. 1985. Theoretical considerations in the measurement of the English proficiency of limited English proficient students. Paper presented at the annual meetings of the National Council on Measurement in Education, Chicago, April.
- Canale, M., and M. Swain. 1980. Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics* 1:1-47.
- Chamot, A.U., and J. M. O'Malley. 1987. The Cognitive Academic Language Learning Approach (CALLA): A bridge to the mainstream. *TESOL Quarterly* 21:227-49.
- Crandall, J., and K. Willetts. 1986. Content-based language instruction: A CLEAR-sponsored seminar. *ERIC/CLL News Bulletin* 9.2.
- Cummins, J. 1982. Tests, achievement, and bilingual students. *Focus* 9, February. Wheaton, Md.: National Clearinghouse for Bilingual Education.

244 / Georgetown University Round Table on Languages and Linguistics 1989

- Darling-Hammond, L., and A. E. Wise. 1985. Beyond standardization: State standards and school improvement. *Elementary School Journal*, January. 315-36.
- Education Commission of the States. 1985. State activity--minimum competency testing. *Clearinghouse Notes*. Denver, Colo.: Education Commission of the States.
- Ellwein, M.C., G. V. Glass, and M. L. Smith. 1988. Standards of competence: Propositions on the nature of testing reforms. *Educational Researcher* 17:4-9.
- Freeman, D., and V. Freeman. 1988. Whole language content lessons. *Elementary ESOL Education News* 1:11.
- Gagne, E.D. 1986. The cognitive psychology of school learning. Boston, Mass.: Little, Brown.
- Gross, S., and J. Stevenson. 1986. Use of a computerized adaptive testing module for ESOL/bilingual entry/exit decision-making: A preliminary study. Paper presented at the annual meetings of the American Educational Research Association. San Francisco.
- Henning, G. 1987. A guide to language testing. Cambridge, Mass.: Newbury.
- Hirsch, E.D., Jr. 1989. The primal scene of education. *The New York Review of Books* 36 (3), 29-35.
- Jacob, E., and B. Mattson. 1987. Cooperative learning with limited English proficient students. Q&A. Washington, D.C.: ERIC Clearinghouse on Language and Linguistics. September.
- Jones, B.F., A. S. Palincsar, D. S. Ogle, and E. G. Carr. 1987. Strategic teaching and learning: Cognitive instruction in the content areas. Alexandria, Va.: Association for Supervision and Curriculum Development.
- Maculaits, J.D. 1982. The Maculaits assessment program. Hayward, Calif.: Alemany Press.
- Martinez, R., and B. Cunningham. 1989. The development of criterion-referenced standardized tests in the content areas for bilingual/ESL student populations. Paper presented at the annual meetings of the American Educational Research Association, San Francisco, April.
- Nelson, J. 1988. Secondary integrated language tasks. Prince Georges County, Md.: Prince Georges County Public Schools.
- Oller, J.W., Jr. 1979. Language tests at school. London: Longman.
- O'Malley, J.M. 1988. Using evaluation to improve instructional services for limited English proficient students. *Forum*, 114.
- O'Malley, J.M., and A. U. Chamot. 1989. Learning strategies in second language acquisition. New York: Cambridge University Press.
- Richard-Amato, P.A., ed. 1988. Making it happen. New York: Longman.
- Saville-Troike, M. 1989. Academic language proficiency: Implications for assessment. Paper presented at the Seminar on Assessing Students in Integrated Language and Content Instruction. Washington, D.C.: Center for Language Education and Research.
- Wise, A. 1988. Legislated learning revisited. *Phi Delta Kappan*. 69. 328-332.

'Passages': Life, the universe, and language proficiency assessment

Thea C. Bruhn
Foreign Service Institute

The debate surrounding the concepts and practices of language proficiency assessment, notably the ACTFL Guidelines and the form of the oral proficiency interview (OPI) used and promoted by ACTFL, continues with few expectations of finding satisfactory resolution before further angst and ample empirical research. What we have not carefully done, however, is to attempt to take issues raised by this debate and look specifically at how they are reflected, or not reflected, in specific cases of application of language proficiency testing. What I shall do in this paper is first, highlight a few of the points that have been raised in the ongoing discussion and then look at how they are 'played out' in two very different cases of usage of language proficiency measures. Finally, in keeping in the spirit of this Georgetown University Round Table theme--lessons from the past with a view toward the future--I will propose a view for attempting to reconcile the issues with 'real life'.

The most damning statement made of the commonly used proficiency rating scales is that they are replete with 'fatal flaws' which render them invalid. I shall not reproduce the complete argument here as it has been amply covered in the literature. (Bachman 1987, Bachman and Savignon 1986, Lantolf and Frawley 1985, Savignon 1985, Barnwell 1989, Freed 1989, Jarvis 1986, Kramsch 1986, Wesche 1983). In essence, the claim is that these rating scales are not empirically based; they have no construct validity nor content validity. They 'are experientially based only in the sense that they are based impressionistically on the experiences of those who drew them up' (Barnwell 1989:44). The scales are further criticized in their use of a native-speaker norm: (1) no native speakers appear to have been involved in the design of the scales; (2) the use of a native speaker as the norm is troublesome due to the difficulty (impossibility?) of defining what a 'native speaker' is; (3) the norm of 'well-educated native speaker' is elitist.

The vehicle used to assign a rating from these scales, the oral proficiency interview has also faced sharp criticism. Bachman has pointed out that a single global rating or holistic score confounds trait and method and leads to a test with no validity. Research suggests that there is no single unitary language ability or 'global language competence' (Bachman and Palmer 1983, Carroll 1983). Ultimately, it is argued that the rating assigned depends more on the method used to collect the sample of language in an individual test situation than on the actual proficiency of the person being rated. The OPI

is essentially a 'one-item' type of test even in the role-play situation. The sample it provides is therefore rather limited even in the role play exercise. Though Byrnes (1987) maintains that a skilled interviewer can elicit a broad rich sample, the fact that a less skilled interviewer may elicit only a narrow limited sample indeed suggests that the examinee's rating will be affected as a result. Although the interview format is intended to be flexible, it

does not allow for meaning to be negotiated, for genuine interaction to take place or for language to be used in a creative, as opposed to a conformist manner.... Finally, it has been suggested that the interactions that are produced by interviews may not be as neutral and natural as was once supposed. (Skehan 1988:217)

Such procedures may be less threatening when the examinee and interviewers share the same background, but can be anxiety-ridden when this is not the case. On the other hand, Canale (1988) points to the tendency of the oral interview to overadapt to speakers' (interviewees') discourse level and cultural patterns.

A communicative test then, such as the OPI, should assess language used for a purpose beyond itself; it has to propose a language-using purpose which can be fulfilled. As Wesche (1983:45) points out, 'the format, content, and scoring criteria of any test should reflect its *objectives*'.

Linguistics has yet to adequately explain how language is systematically related to experience, so we are not always certain exactly which variables in communicative acts are the important ones. Still, it would seem that at least the following information should be clarified for each communicative act in the objectives: the *purpose* of the interaction (including which topics will be treated, related notions, and the language functions which the learner will need), *situational* aspects which will influence language behavior (including the social and psychological roles and relationships of the participants and the settings in which communicative interaction will take place), and the types of *discourse* which will be appropriate (re: genre, variety, visual or auditory channel, etc.). It is also important to determine the *degree of skill* expected of the learner or examinee. Such definition of objectives will make it possible to determine the *language forms* (structures, words and phrases) needed by the second-language speaker, or at least to specify the *kinds of authentic materials and interactions with native speakers* which would expose the learner or examinee to appropriate forms. (Wesche 1983:45)

Against the background of this discussion, let us take a look at two different cases of application of language proficiency testing. The first case is that of the Foreign Service Institute (FSI); the second is the Summer Institute of Linguistics (SIL) Bilingualism Project in Senegal, West Africa.

The FSI originated the descriptions and methods upon which current proficiency guidelines and current oral proficiency assessments are based. The development of the so-called FSI guidelines and FSI test was promoted in the 1950s initially in response to the National Mobilization and Manpower Act,

which directed the Civil Service Commission to inventory and develop a register of persons in government with foreign language skills. It was not until 1979, in fact, that FSI undertook a project to promote the use of its testing system outside of the federal government (Frith 1979).

The original goal of the system was to provide survey data of foreign language speakers in the foreign service. It began as a self-rating based on six language user descriptions. The lack of differentiation of skills or components has been attributed to the desire for simplicity. It played into prevailing views of psychometrics which spent a lot of time and effort in the first part of this century to demonstrate the existence of a single measurable attribute called general intelligence--the 'G' factor. Now I can only conjecture that this concept may have provided the authors of the original language proficiency level descriptions with a handy framework for assuming an 'L' factor or unitary language proficiency that evolved into the 'Global Rating.' Another assumption which appears to have carried over from intelligence testing was 'that the language proficiency of an individual is at a single point on a linear scale, and that this point can be determined by an objective test instrument' (Underhill 1987:5).

The self-rating proved to be unreliable and the use of a single rating unworkable. FSI decided then that the rating would be made by a 'scientific linguist' on the basis of an observed conversation between the subject employee and a native speaker of the language and that each person would be rated separately in speaking and reading. A checklist for speaking was developed that included five factors--accent, grammar, vocabulary, fluency, comprehension--with each factor divided on a linear six-point scale. The purpose of the factors and checklist was to balance the inherent subjectivity of the testing procedure by providing agreement about what aspects of the performance were to be observed and how. A weighted system for scoring the checklist was subsequently developed (Wilds 1965). The face-to-face conversation evolved into a structured interview with each testing team developing its own 'tricks of the trade' for zeroing in on what it regarded as the salient features of the language based on the factors and existing guidelines.

This by and large is the system which was adapted and adopted by other federal government agencies and was eventually packaged for export outside the government.

The reporting requirements of the 'Uniform Guidelines on Employee Selection Procedures' under Title VII of the Civil Rights Act, issued in 1978, appear to have prompted FSI to conduct several validation and reliability studies in an effort to support empirically both the FSI Skill Level Definitions and the testing procedure. FSI concluded that its testing system was valid and reliable. Yet even as these assertions were made and the system was being promoted to academia, internal memos reveal widespread dissatisfaction, particularly among the 'consumers,' i.e. the foreign service officers subject to the system (Adams 1981?).

The system was unable to make meaningful judgments about or to measure effective communication. The existing assumption was that a face-to-face conversation was a valid replication of professional function, which it is not. A person's so-called language proficiency, while it may have been quite

accurate in terms of technical skill, did not mean effectiveness in communication. The native speaker performance standard presumed all examinees would perform at level 5 (highest) if tested in their native language, but which native speakers of what language variety did this refer to?

There were no precise criteria for what constituted a good sample. Informal general lists of features allowed too much variation from test to test. Along with this came the subjective nature of scoring. Even Wilds (1975) points out that the rating each examiner would give was normally not based on the checklist, and the actual determination of the S rating was handled differently from team to team. A survey of FSI testing practices conducted in the late 1970s and early 1980s uncovered a wide diversity in technique and judgment.

A revamped FSI training curriculum in 1981 brought additional pressure upon the testing system. The sole function of the FSI test, it was proposed, should be to forecast, by reference to an agreed-upon set of 'functional linguistic criteria,' the expectation that a post may have of an officer's ability to use the foreign language for the accomplishment of foreign-affairs relevant tasks. The criteria in the existing test had been derived from a particular structuralist analysis of Indo-European languages. There was a perceived need to develop a system from an independent sociolinguistic analysis of situated language usage in foreign-affairs contexts in post countries. There was a need for the FSI test to be pragmatically not academically oriented. Also recognized was the need to maintain a consistency of viewpoint about communicative competence across languages, but to abandon the notion of strict comparability (i.e. a '3' knows the subjunctive, a '2' uses the past, present and future tenses, etc.).

Essentially, FSI threw out the 'FSI test.' A conscious decision was made, however, to do so quietly. Wider applicability of FSI testing procedures either in the academic world or in the interagency context was made a low organizational priority until successful results were obtained on the synchronization of the testing system and the revamped FSI curriculum. A curriculum which was pragmatically and communicatively oriented could not be fully implemented, it was felt, until the testing system reflected similar orientation and goals, i.e. effective communication.

FSI developed a new test with a structured format and standard stimuli in an attempt to ensure comparability among languages and obtain a richer, more representative sample of language from the examinee. The titles of three separate parts--conversation, interview, briefing--reflect representative professional functions identified in an analysis of foreign service officer language use at post. Revised factors were developed along with generic standards of performance which serve to identify aspects of an examinee's performance that contribute to, or detract from, communicative effectiveness.

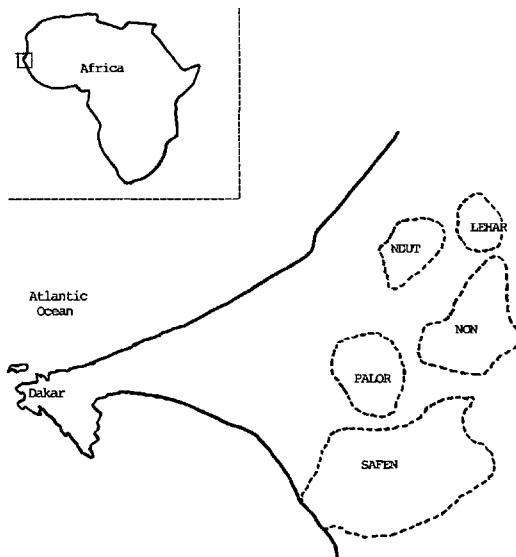
This second generation FSI test then is designed from a needs orientation rather than a language orientation. It is not governed by an *a priori* set of assumptions about speech events and language but assumes a practical orientation toward testing language proficiency based on observations of post language use. The fit of tested scores and observed proficiency has not been empirically determined but is based on information which FSI considers to be more than anecdotal. The current FSI approach reflects a philosophy of

proficiency that does not imply universal or global proficiency, but rather limited proficiency. It does not make rating dependent on an implicational scale and is not bound by a theory of what is high or low level language use, but by observation; there are certain things '3s' can do and '2s' cannot. The present FSI test fills an observed need. The attitude is essentially that if the theory does not support what FSI is doing, then there is probably something wrong with the theory.

The FSI Test II (to distinguish it from the original FSI oral interview) addresses and remedies many of the problems experienced with the OPI. There remain issues, however, which it continues to ignore or fudge over. The ILR Language Proficiency Skill Level Descriptions, the present-day version of the FSI guidelines, maintain an educated native-speaker norm. Despite the use of a scoring procedure which attempts to circumvent the issue of the native speaker as yardstick, the FSI Test II still displays this inherent bias. The troubling questions remain: when people who are not 'well educated' are evaluated, do the ratings reflect their ability to use the language or their lack of education; how does their performance compare to speakers of the language of a comparable education, occupation, or socioeconomic level?

Our second case drives home this thorny issue even further and reinforces the problematic nature of claims of generic proficiency. In 1986, I was contacted by the Summer Institute of Linguistics (SIL) to work with their Bilingualism Project to assist them in developing methods of assessing the degree of bilingualism of speakers in communities identified by language surveys as being bi- or multilingual. The initial work would be done in West Africa among speakers of five languages/dialects in the Cangin complex: Non, Ndut, Palor, Lehar, and Safen (Figure 1). The language to be assessed initially would be Wolof, a lingua franca throughout much of that area of Africa.

Figure 1. Distribution of Cangin languages/dialects in the westernmost region of Senegal.



Among other things, the Bilingualism Project attempts to evaluate a language population by sampling to find out if most of a population can use a second language well, without domain restriction, for all communication needs and situations. Further, it studies the language group's attitudes toward the second language with relation to ethnic identity and interethnic relations. Information such as this is essential for local or national governments in multilingual areas to develop reasonable and coherent policies on education, literacy development, and language planning which have any chance of success.

SIL has a contract with the Literacy Development Office of the Education Ministry of the government of Senegal to do linguistic surveying in Senegal and assist the government with literacy development. Because of the large number of languages spoken in Senegal, it would be impractical and costly to attempt to develop literacy and education in each language. Furthermore, such linguistic fragmentation of a population would work against the goal of national unity which most governments have. More efficient and desirable is the possibility of identifying a few widely used languages to use in literacy and education development. To do this, it is necessary to assess the extent to which languages of wide communication, lingua francas, are known by the population. If people are illiterate but proficient in the second language, they can be taught to read and can learn in that language. Having to teach them a second language in order for them to learn to read and receive instruction is a much more difficult and complex task.

In undertaking the bilingualism survey, SIL realized that it lacked an effective means for assessing the level of proficiency that a given group may possess in a second language. A careful review by SIL's international survey coordinator, Dr. Ted Bergman, of existing approaches to assessing language proficiency led him to conclude that the approach which FSI currently uses would most closely meet their needs and would have the best potential for being able to be adapted to the peculiarities of minority language group situations.

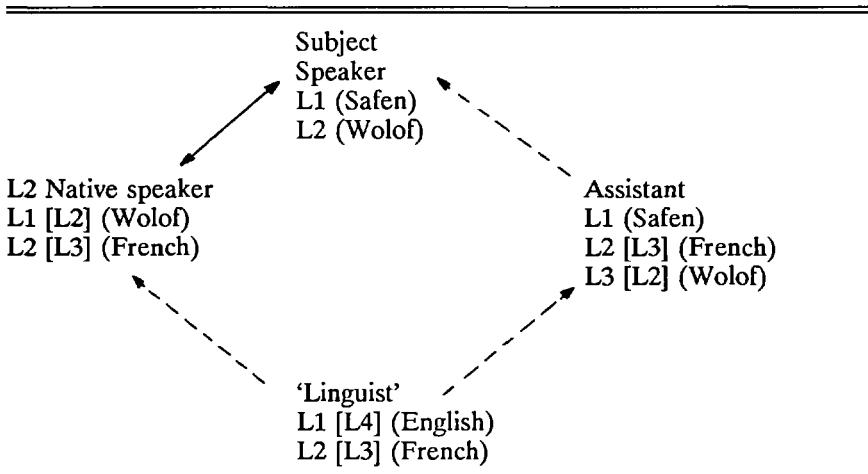
The first few days of the phase of the project in which I was involved were spent doing a thorough needs assessment to determine what the specific objectives and constraints would be for a language proficiency assessment instrument in the linguistic survey context. This entailed a review of related survey efforts, a profile of typical survey populations and conditions, and a discussion of related research on testing of language proficiency. Of primary concern was the fact that the majority of people who would be tested are non-literate, noneducated, and lacking in familiarity with westernized culture and technology. The former characteristics impact on the specific techniques FSI uses to collect a language sample in the course of a test, the latter on the venue for a test.

Time was spent discussing the principles underlying FSI's approach to testing language proficiency. These were then examined from the perspective of the field conditions and population characteristics with which SIL would be working in Africa as well as other parts of the world. It was determined that the descriptions for skill levels were adaptable to the SIL goals and objectives. Likewise, a tripartite oral assessment and the evaluation procedure such as FSI employs were deemed workable with modifications. The revised

procedure became a conversation, followed by a task requiring the subject to get information, followed by a task in which the subject relates information.

Modifications to the standard FSI procedures and descriptions were drafted and tried out on some local volunteers. The most significant difference in the procedure is the introduction of a third member to the testing team, an Assistant. The FSI test relies heavily on written instructions and information to use with examinees. This of course is not feasible with non-literate populations. The Assistant is a bilingual speaker who has as mother tongue the same language as the subject, and a second language with which to communicate with the person administering the test. The Assistant takes on the role of the written instructions and information in the FSI test, that is, provides the subject with instructions, with suggestions, and with information in the subject's mother tongue.

Figure 2. Testing team configuration and language/participant interaction in SIL proficiency assessment.



A village named Paki, situated approximately forty miles southeast of Dakar, was selected as a site for piloting the testing procedure. The inhabitants of Paki speak Safen, but are bilingual in Wolof. Wolof was selected as the language in which to conduct the trial tests.

The initial trial tests pointed to the need to develop some specific guidelines for the conduct of tests as well as for the types of materials or topics used during the test. Due to cultural differences, there are several issues which arose that FSI does not face. Sex, age, and socioeconomic factors enter the assessment situations. The use of a higher status male as a tester with a lower status female produced a skewed sample. The use of a younger male or female tester with an older male or female subject also produced sampling difficulties as well as evaluation problems. The younger person was unwilling in any way to criticize or observe shortcomings in the older person's use of language, no matter how incorrect.

Figure 3. A comparison of the characteristics of the application of the FSI and SIL assessments.

FSI	SIL
Setting: Office/interior room at FSI or overseas post	Remote village
Participants: Educated English-speaking FSO's Speaker 'Linguist'	Nonliterate speakers of various L2 native languages/dialects L2 native speaker 'Linguist' L2/L3 assistant when 'linguist' does not have speaker L1
Goals: Make personnel decisions on basis of ratings (of individual)	Make policy decisions on basis of ratings (of group)
Format: 3 parts: conversation, interview, briefing (occupational relationship)	3 parts: conversation getting information, relating information (language community relationship)
Interreactional norms: Written instructions in L1 outlining process, expecta- tions	Oral instructions in L1 outlining process, expectations
Language context: English L2 (various)	L1 (variations) L2 (L3)

The power-solidarity semantic has not gone unnoticed in other applications of oral proficiency testing. As Byrnes (1987:173) points out:

There can be no doubt that the interview gives primary evidence for only one kind of oral language use, namely the response a second-language speaker is able to provide to a single individual, the tester, in a setting that is marked by a power relationship....

Even more interesting was the problem of trying to determine guidelines for what constitutes the standard form of the language, especially in the

instance of a language such as Wolof which exists along a continuum. A pragmatic approach was outlined and stated in the drafted SIL Guidelines.

In situations where education does not necessarily define the best speakers of the standard variety of the target language, the variety used by any existing media (e.g. radio, television, etc.) should be reflected in determining the standard. Respected people in the [target language] community should be consulted in determining the standard variety of the target language...

What we see then are two very different cases of the use of a similar oral assessment technique. There are marked differences in the setting, participants, goals, and interactional norms and language contexts between the FSI and SIL assessments which impact on the actual format, sampling and evaluation.

From these examples we can see evidence that it is not possible to have a universally applicable rating scale or rating system. As has been pointed out elsewhere, tests which are more reflective of the kinds of situations, language context, and purposes for which L2 speakers need their skills, enable us to make more accurate predictions about how well they will be able to function using the L2 in 'real life' (Wesche 1983:53). Of continuing concern is the effect of the method and sociolinguistic variables on rating outcome, even assuming rating criteria based on need, and language use data relating to communicative effectiveness. There is certainly anecdotal evidence in the case of FSI that this is a very real and present problem, and a very stark example in the case of the SIL Bilingualism Project.

The SIL has been piloting the oral assessment procedure for a year now. It would be interesting to look at how the ratings it produces compare to the survey data and other measures of bilingualism that are being applied.

Examining particular cases such as these applications of proficiency measures raises doubts about whether the elusive 'common metric,' a standard for defining and measuring language proficiency independent of specific languages, contexts and domains of discourse, is a feasible or even desirable goal.

Note

I am indebted to H. David Argoff, Associate Dean of the School of Languages, Foreign Service Institute. He has served as an 'institutional memory' for me, providing information and insight to fill the gaps in the written literature on the FSI test. Any interpretation or misinterpretation of that information, however, remains solely my own.

References

- Adams, Marianne L. 1981 (?). Validation and reliability of the FSI oral interview. Unpublished papers and memos.

254 / Georgetown University Round Table on Languages and Linguistics 1989

- Bachman, Lyle F. 1987. Problems in examining the validity of the ACTFL oral proficiency interview. In: *Proceedings of the Symposium on the Evaluation of Foreign Language Proficiency*, ed. Albert Valdman. Bloomington: Indiana University Press. 29-43.
- Bachman, Lyle F., and Adrian S. Palmer. 1983. The construct validity of the FSI oral interview. In: *Issues in language testing research*, ed. John W. Oller. Rowley, Mass.: Newbury House. 154-69.
- Bachman, Lyle F., and Adrian S. Palmer. 1979. Convergent and discriminant validation of oral language proficiency tests. In: *Proceedings of the Third International Conference on Frontiers in Language Proficiency and Dominance Testing*, ed. Raymond Silverstein. Carbondale, Ill.: Southern Illinois University. 53-62.
- Bachman, Lyle F., and Sandra J. Savignon. 1986. The evaluation of communicative language proficiency: A critique of the ACTFL oral interview. *Modern Language Journal* 70.4:380-390.
- Barnwell, David. 1989. Proficiency and the native speaker. *ADFL Bulletin* 20.2:42-46.
- Byrnes, Heidi. 1987. Features of pragmatic and sociolinguistic competence in the oral proficiency interview. In: *Proceedings of the Symposium on the Evaluation of Foreign Language Proficiency*, ed. Albert Valdman. Bloomington: Indiana University. 167-77.
- Canale, Michael. 1988. The content validity of some oral interview procedures: An analysis of communicative problems and strategies. Paper presented to the 10th Language Testing Research Colloquium.
- Carroll, John B. 1983. Psychometric theory and language testing. In: *Issues in language testing research*, ed. John W. Oller. Rowley, Mass.: Newbury House. 80-107.
- Freed, Barbara F. 1989. Perspectives on the future of proficiency-based teaching and testing. *ADFL Bulletin* 20.2:52-56.
- Frith, James R. 1979. Testing the FSI testing kit. *ADFL Bulletin* 11.2:12-14.
- Jarvis, Gilbert A. 1986. Proficiency testing: A matter of false hopes? *ADFL Bulletin* 18.1:20-21.
- Kramsch, Claire J. 1986. Proficiency versus achievement: Reflections on the proficiency movement. *ADFL Bulletin* 18.1:22-24.
- Lantolf, James P., and William Frawley. 1985. Oral proficiency testing: A critical analysis. *Modern Language Journal* 69.4:337-45.
- Savignon, Sandra J. 1985. Evaluation of communicative competence: The ACTFL proficiency guidelines. *The Modern Language Journal* 69.2:129-34.
- Skehan, Peter. 1988. Language testing. *Language Teaching* 21.4:211-21.
- Underhill, Nic. 1987. Testing spoken language: A handbook of oral testing techniques. Cambridge: Cambridge University Press.
- Wesche, Marjorie Bingham. 1983. Communicative testing in a second language. *Modern Language Journal* 67.1:41-55.
- Wilds, Claudia P. 1975. The oral interview test. In: *Testing language proficiency*, ed. Randall L. Jones and Bernard Spolsky. Arlington, Va.: Center for Applied Linguistics. 29-44.

Language testing in the secondary schools: Past experience and new directions

Rebecca M. Valette
Boston College

Over thirty years have gone by since the Russian launching of Sputnik and the subsequent National Defense Education Act which gave a strong impetus to foreign language instruction in American schools by helping finance the creation of new teaching materials (the A-LM series or 'audio-lingual method') and new testing batteries (the MLA Coop Tests). Language laboratories sprang up across the country, and hundreds of teachers were able to strengthen their foreign language skills at NDEA Institutes both here and abroad. It is now appropriate to look back over the past three decades to determine to what extent language testing techniques and procedures in the secondary schools have evolved and what new directions might profitably be explored.

The main focus in this paper is on foreign language instruction at the secondary school level in the United States. Roughly 15-20% of American high school students study a second language. Of this group, almost two-thirds study Spanish, less than one-third study French, and the remainder are enrolled in German, Italian, Russian, with a sprinkling in the less commonly taught languages. Most of these secondary school language students will not travel abroad or use the new language for real communication: for them, the language course is simply another subject in the curriculum, one recommended or required for college entrance. (An exception to this generalization are the Hispanic-American students enrolled in Spanish courses.)

Statistics as to how much language these students are exposed to in the course of their study have not been recently established, but I would venture to guess that one-half to two-thirds of all students enrolled in a given language sequence do not go much beyond 'Level One', if we define this level as consisting of the material contained in the first book of a standard secondary school language program. (For French and Spanish students, if we use traditional 'grammar' terminology, this would consist of learning to handle the present tense and the passé composé or preterite with some notions of the imperfect.) On the ACTFL oral proficiency scale, most of these students are probably at the Novice Level. One-fourth to one-third of the secondary school language students probably complete 'Level Two' (in which they consolidate material from Level One and are introduced to reflexive verbs, other indicative tenses, more pronouns and some basic notions about the subjunctive). On the ACTFL oral proficiency scale, these students are probably for the most part at Novice High, with some at

Intermediate Low. Only a very small percentage of the secondary school language students complete the 'Level Three' or higher.

These assessments are 'guess-timates' as to the national picture. Obviously, much higher percentages of students are enrolled in language courses and follow longer sequences of instruction in the affluent suburbs, while in certain other parts of the United States the percentages are much lower and language instruction, where offered, may not go beyond a 'Level One' program spread over two years. Similarly, teachers at the affluent schools are likely to exhibit an acceptable to high degree of oral proficiency in the foreign language and will tend to encourage the development of oral proficiency in the classroom. Elsewhere foreign languages may be taught by teachers who feel uncomfortable speaking the language and who prefer to focus on the less threatening and more familiar areas of grammar and vocabulary learning.

Against this backdrop of secondary school language instruction in the United States, how has language testing evolved over the past thirty years and what role does it play in the curriculum? Here, again, there has been no national survey. What follows is the product of informal contacts with teachers around the country and the result of three decades of professional interest in the topic.

Developments in language testing. In the 1950s, the focus of secondary foreign language instruction was on grammar, vocabulary acquisition, and reading. Typically, English was the frame of reference and the language of the classroom. The students enrolled in language classes were generally college-bound and constituted the academic elite. Classroom tests which accompanied this 'traditional' cognitive type of language instruction usually contained the following types of items:

- a. Vocabulary lists, usually from English to the foreign language;
- b. Incomplete sentences, for which the student would provide the equivalent of an English cue or the appropriate form of a verb given in the infinitive form;
- c. English sentences to be translated into the foreign language;
- d. Foreign language passages to be translated into English;
- e. Dictations.

In the 1960s, the audiolingual movement, with its close ties to structural linguistics and behaviorism, encouraged teachers to focus on the four language skills of listening, speaking, reading, and writing. This 'New Key' approach was accompanied by a new view of language testing based on certain key tenets:

- a. Translation was to be avoided;
- b. Students were not to engage in 'code switching', that is, test items would not mix English and the target language;
- c. 'Pure' tests, that is, tests which focused on one language skill at a time, were preferred over 'hybrid' tests which mixed language skills;
- d. Test items would not contain wrong linguistic forms.

Clearly, the traditional language testing formats were no longer considered valid. Translation was out. Dictations which mixed listening and writing skills were not 'pure' tests and were to be avoided. Sentences that mixed two languages were found to be linguistically unsound and therefore were banned. Only items that cued a verb form by its infinitive could still be included in the new language tests.

In the 1960s, two sets of standardized tests were developed in modern languages: the MLA Cooperative Foreign Language Tests (1963) for secondary school and college students, and the Pimsleur Proficiency Tests (1967) for secondary school students. Both of these batteries, reflecting the new emphasis on four skills, consisted of four independent sections labeled Listening, Speaking, Reading, and Writing. It is informative to examine the types of items used in these tests, for their format influenced classroom testing.

Let us review the Pimsleur Spanish Proficiency Test, which was developed in two forms: Form A, for administration after a 'first-level' course, and Form C, for administration after a 'second-level' course.¹

Listening Test

Part 1. Matching a spoken sentence with its written equivalent.

tape: El tiene razón.

- booklet:
- a. Ella tiene razón.
 - b. Ellas tienen razón.
 - c. El tiene razón.
 - d. Ellos tienen razón.

Part 2. Selecting the appropriate response to a spoken question.

tape: ¿Qué tiempo hace?

- booklet:
- a. Es el dos de mayo.
 - b. Hace frío.
 - c. Son las seis.
 - d. Está bien.

Speaking Test

Part 1. Identifying nouns

picture: loaf of bread

correct response: pan

Part 2. Reading aloud

cue: Esta chica está sana.

correct response: appropriate syllabic stress on *esta* and *está*

Part 3. Answering questions

tape: ¿En qué mes nació Vd.?

- scoring: 3 points--complete sentence, perfect content and structure
 2 points--complete sentence, appropriate and understandable in spite of error(s)

1 point--meaningful response: incomplete sentence or poorly structured or slightly inappropriate

Reading

Short reading passages followed by multiple choice questions

Writing

Part 1. Sentence completion

Write the *single* missing word.

El árbol está cerca de la casa.

Part 2. Verb forms

Form A: Complete the sentence with the correct form of the verb given in parentheses. Use the same tense as the other verb in the sentence given in darker type.

Juan **baila** y canta (**cantar**) bien.

Form C: Complete the second sentence with the appropriate form of the missing verb.

Me dijeron que iban a hacerlo.

Me dicen que van a hacerlo ahora.

Part 3. Sentence transformation

Rewrite the sentence replacing the underlined word with the word in parentheses and making all other necessary changes.

Veo el edificio bonito. (**casa**)

Veo la casa bonita.

Part 4. Describing pictures

Form A: Describe each picture with a sentence. Some take place in the present and others in the past.

picture cue: male teacher talking to male student ('ahora')

response: El maestro habla al alumno.

scoring: 3 points: excellent sentence (one wrong accent mark or capitalization error allowed)

2 points: fairly well written, correct time perspective, verb may be misspelled

1 point: well-written sentence that misses the point or poorly written sentence that communicates the main action

Form C: Write a description of what Carlos and Pedro did yesterday.

picture cue: sequence of 8 frames: Pedro arrives at Carlos' house and they study together; the mother serves hot chocolate; Pedro leaves.

scoring: 1/3 completeness of story and choice of vocabulary

1/3 correct verb formation and usage

1/3 correct spelling and grammar (other than verbs)

What is striking to us today as we look back twenty years is that although the tests are labelled 'listening', 'speaking', 'reading', and 'writing', these terms tend to refer simply to the basic format of the items. The main emphasis of

the test is on linguistic accuracy: pronunciation, spelling, grammar (especially verb forms). Moreover, those sections which focus on meaning often resemble discrete vocabulary tests (naming pictures, knowing the meaning of an expression in the reading passage). There is little place for natural communication. Even in the 'interview section' (part 3) of the speaking test, full sentence answers are required.

Classroom tests in the 1960s tended to follow a model similar to that of the new standardized tests. Oral grades were based on the recitation of dialogs, with the emphasis on accuracy of pronunciation and no attention given to meaning. Written tests were of the fill-in or transformation variety, again with a focus on correct form. (In fact, the teachers' preoccupation with accurate spelling and form was reflected in the new generation of 'revised audiolingual' textbooks that contained more grammar charts and vocabulary lists.)

It is instructive to note that the students of the 1950s had to know the meaning of what they were writing as they translated:

'I looked for my aunt's pen.'
J'ai cherché la plume de ma tante.

In contrast, the students of the 1960s and 1970s when faced with the following item needed only to focus on form:

Je cherche la plume de ma tante.
Hier j'ai cherché la plume de ma tante.

The student might reason as follows: 'Hier' signals that I need the passé composé; this means that I write 'j'ai' and then put an acute accent to form 'cherché'.²

In the 1970s, there was a broadly publicized effort to make language learning more accessible to students of various levels of aptitude. In 'individualized instruction' the desired learning outcomes of each lesson were described in terms of behavioral objectives, and students were expected to attain a certain score on the corresponding lesson tests before moving on to the next unit. This meant they had the opportunity to take a test several times if necessary in order to obtain a passing grade.

One of the behavioral objectives of a first-year French lesson might be the following:

Objective: To form the passé composé of regular -ar verbs.
 Students must complete 80% of the sentences correctly.
 1. Hier Marie _____ la plume de sa tante. (chercher)
 2. Nous _____ nos livres. (trouver) Etc.

Although all types of course objectives could be stated in behavioral terms, in reality most teachers focused on grammar and pronunciation objectives that could be rapidly tested and readily scored. Form was more important than meaning, and communication goals were often given little attention. A serious problem faced by many individualized programs was the

increased attention paid to the written language at the expense of providing practice using the spoken language.

The mid-1970s saw the publication of *French for Mastery*, the first American secondary school textbook series to incorporate the research of the Council of Europe into second language learning and the elaboration of the functional-notional syllabus. The focus was on meaning and all activities were presented in a situational context. The key concern was communication: understanding what one hears and reads, and expressing one's ideas in speech and writing. Units were organized around daily-life themes and goals that were expressed in functional terms. For example, chapter 5 focused on food. The students were told:³

You will learn how:

- to talk about the foods you like
- to read a French menu and order a meal in a French restaurant
- to prepare a shopping list
- to make suggestions to your friends

The vocabulary and grammar of the chapter were derived from the theme: names of foods and beverages, use of the partitive article, use of the imperative. The program was 'grammar-supported' rather than 'grammar-driven'.

The corresponding testing program did reflect the secondary school emphasis on teaching correct forms. Students were expected to demonstrate their mastery of the new vocabulary and structures; however, test items were presented in situational contexts. For example,⁴

Au club de vacances. It's buffet night. Claire is taking the things she likes and not taking things she does not like.

Elle aime la soupe. Elle prend de la soupe.

Elle déteste le fromage. Elle ne prend pas de fromage.

The classroom achievement tests accompanying the program did also include communicative listening test portions entitled 'Messages' where students demonstrated their comprehension by taking notes. For example, in the test corresponding to the foods chapter, students were asked to imagine they were staying with a French family and were to write up the shopping list as explained by their hostess.

It was not until the mid-1980s that broad segments of the foreign language profession began to express a commitment to teaching toward communication goals. The promulgation of the ACTFL Proficiency Guidelines involved teachers and groups across the country. Most of the new secondary school textbook programs have now become grammar-supported, rather than grammar-driven, and emphasize meaningful language use.

Several states, including Wisconsin and New York, have produced new curriculum guides emphasizing proficiency. The State of Minnesota established university entrance and exit requirements based on proficiency standards.

In 1988, the New York State Board of Regents issued guidelines for French, German, Italian, and Spanish proficiency examinations which would be available for administration in Fall 1989. At 'Checkpoint A', intended primarily for eighth grade students who have completed two units of instruction, these examinations would consist of four parts:⁵

Part 1. Speaking (30%)

1a. Informal classroom evaluation (10%)

1b. Formal speaking test (20%)

four short conversational tasks, one in each of the following categories:

A. Socializing

B. Providing and obtaining information

C. Expressing personal feelings

D. Getting others to adopt a course of action.

Example: *Providing/obtaining information*

[student initiates] Teacher says: You are in a store, looking at clothes. I am a salesclerk. You are going to tell me what you are looking for. You will start the conversation.'

Scoring: Teacher may make two attempts at eliciting each of 4 student utterances; each comprehensible and appropriate utterance is worth 1 point. A fifth point is given for overall quality. Maximum score for each task: 5 points.

Part 2. Listening comprehension (30%)

For each subsection the students hear several short spoken passages read by the teacher from a script. The setting is given in English, and then each passage is read twice. Students mark a corresponding multiple-choice item in their test booklet.

2a. Multiple-choice responses in English

2b. Multiple-choice responses in target language

2c. Multiple-choice responses among four pictures

Part 3. Reading comprehension (20%)

Students look at pieces of authentic realia (ads, forms, short newspaper articles) and answer a general multiple-choice item on each one.

3a. Multiple-choice responses in English

3b. Multiple-choice responses in target language

Part 4. Writing (10%)

4a. Notes (6%)

Students write two notes of at least 12 words each.

Example: 'Your friend has invited you to her birthday party. Write her a note in [language] to explain why you cannot come to her birthday party.'

4b. Lists (4%)

Students write 2 lists of 4 items each.

Example: 'Your pen pal from [country] is coming to spend a few weeks at your house. In [language], write a list of 4 things you will suggest you can do together.'

From a measurement point of view, a major weakness of the New York State proficiency examinations, as described, is the role of the classroom teacher in administering the Speaking and Listening tests. First, this lowers the 'reliability of the speaking portion', since some teachers may be more effective than others as conversation partners in eliciting and encouraging student responses. Moreover, in scoring that portion, some teachers may be more generous than others in assigning the corresponding points. Second, it lowers the 'reliability of the listening portion', since some teachers may read the script more slowly and distinctly than others.

Finally, the statewide use of the voices of so many classroom teachers, some native and many not native, lowers the 'validity of the listening portion' (and to a lesser extent the validity of the speaking portion). As a result, most students will not be listening to authentic speech. This problem of validity could be readily solved by putting this part of the examination on cassette.

In criticizing the proposed format of the New York State Proficiency Examinations, however, we should not forget that the real importance of tests at the secondary school level lies not in their reliability and validity but in their ability to influence instruction. At the classroom level, it is recognized that students study and learn those things on which they are graded and tested. Extramural tests play a similar guiding role in that teachers organize their instruction so as to help their students perform well on these measures. In this sense, the New York State Proficiency Examinations will encourage teachers to provide classroom practice in the types of activities included in those tests. In this sense, it is even more important that recordings rather than teacher voices be used, so that students learn to understand authentic speech.

Looking to the future. What is the situation of the secondary school language teacher at this time as we look ahead to the 1990s? The average teacher feels uneasy. Let us look at several aspects of this problem.

(1) **Insecurity about one's ability in the target language.** Most secondary school teachers in the United States are not Superior, or even Advanced, speakers of the language(s) they teach. They feel insecure listening to authentic radio and video broadcasts because they cannot understand everything they hear. They tend to teach in English because that is the language they command best. Those on the lower end of the proficiency scale find themselves much more comfortable teaching basic grammar structures and giving paper-and-pencil tests than teaching oral communication skills. When they do use the target language in class, they often speak it in a distorted manner.

How can teacher proficiency be improved? Probably one of the most exciting developments in secondary school teaching programs will prove to be the development of integrated video programs filmed abroad and paced to accompany beginning classes. By playing these videos many times in the course of the school day, the weaker teachers will significantly increase their listening skills and gain self-confidence.

(2) Concentration on the written language. Faced with over 100 students a day, most teachers find it expedient to assign written classwork where each question has only one correct answer. In addition, many teachers tend to insist that students be able to write correctly every word or phrase they learn. For instance, when teaching numbers, they want students to learn how to spell them also. (In French does 81, 'quatre-vingt-un', require hyphens? does the 'vingt' take an 's'?) They fail to realize that unless the students have checking accounts abroad, a highly unlikely prospect, they will never be required to write numbers out. Teachers are looking for guidance as to how to balance the spoken and written language in the demanding classroom environment.

(3) Importance of extramural tests. Secondary school teachers want their students to do well on extramural tests. Foremost among these tests is the College Board Achievement Test, which has barely changed over the past thirty years. To obtain a high score, students need a large reading vocabulary and a good control of grammatical structures. Listening skills are not tested at this time, nor are the abilities to express oneself orally or in writing. Teachers preparing students for college entrance feel impelled to spend time teaching grammar and vocabulary. Similarly, the contests sponsored by the professional organizations (e.g. the AATF, AATSP, AATG) use tests that measure primarily grammar and vocabulary knowledge. The fact that certain states are instituting extramural tests which focus on communication skills will pose a potential dilemma as to how these apparently contradictory objectives can be reconciled.

(4) Belief in the importance of linguistic accuracy. Most nonnative teachers have learned the second language in courses where the focus was on formal language use and accurate spelling. Although many have had the opportunity to live or study abroad at some point, they still cling to the notion that the teacher's role is to insist on complete sentences and to correct errors whenever they occur. Even when engaged in the 'communication' activities found in their textbooks, they tend to pay close attention to the form of a student's response and fail to react in a natural manner to what the student has said.

This matter of linguistic accuracy, however, is also of great concern to the teachers committed to developing real proficiency and communicative language use. While it is recognized that a person with high linguistic proficiency (knowledge of grammar, vocabulary, etc.) might be anywhere from low to high on the scale of communicative language use, it is also recognized that a person with low linguistic proficiency will never rate high on the communicative scale.⁶ If the committed teacher wants to lay a solid foundation on which students can eventually progress to a level of high proficiency, is it not important to establish a firm linguistic base?

It will be the challenge of the 1990s to respond to the needs of the secondary school teachers and to create test batteries for Level One (and Level Two) students that reflect the constraints of the classroom situation in the United States and at the same time encourage the building of a strong

target language foundation on which those students so inclined can subsequently develop real proficiency.

Notes

1. *Pimsleur Spanish Proficiency Tests*, New York: Harcourt Brace and World, 1967.
2. See Carol Hosenfeld's work on learning strategies of language learners as reviewed in her article, 'The new student role: Individual differences and implications for instruction', in: Gilbert A. Jarvis, ed., *Perspective: A New Freedom*, Skokie, Ill.: National Textbook Company, 1975, 129-67.
3. Jean-Paul and Rebecca M. Valette, *Workbook for French for Mastery 1*, Lexington, Mass.: D.C. Heath and Co., 1975, 58. It is interesting to note that these statements were deleted from the 1980 edition since the majority of the French teachers across the country were not looking at their courses in terms of developing proficiency.
4. Jean-Paul and Rebecca M. Valette, *French for Mastery 1 Testing Program*, Lexington, Mass.: D.C. Heath and Co., 1976, chapter 5, 3.
5. New York State Proficiency Examination in Modern Foreign Languages: An Introduction. (Ditto.) Albany: State Education Department, Bureau of Foreign Languages Education, October 1988.
6. For more on this topic, see Paula M. Rea Dickins and Edward G. Woods, 'Some Criteria for the Development of Communicative Grammar Tasks', *TESOL Quarterly* 22.4:623-46.

Who is in charge in the learner-curriculum-testing connection?

Heidi Byrnes
Georgetown University

Introduction. It is probably no overstatement to observe that much of the excitement generated in the foreign language teaching community in the last decade is closely associated with the term 'proficiency'. While the term lacks definitional tightness, it is generally taken to refer to the ability to function appropriately in a language in all skill modalities in a wide variety of contexts and social settings. Of necessity this entails grammatical, discourse, pragmatic, and sociolinguistic competencies.

Such a description places proficiency in close proximity to the older term 'communicative competence'. However, the repercussions of the proficiency orientation have been decidedly more practical than those of its antecedent, communicative competence, and have been felt throughout all levels of instruction. This fact is at once curious in light of the proficiency movement's provenience and at the same time not totally unmotivated and inappropriate: Although the term 'proficiency' entered our professional consciousness via the back door of testing oral language use, its greatest impact seems now to be directed toward curriculum, instruction, and materials development, as evidenced by the remarkable number of such activities and elaborate initiatives at all levels of the American educational establishment. The ACTFL Proficiency Guidelines are currently perhaps the most widely cited document for justifying a variety of approaches in methodology, in curriculum and syllabus design, as well as materials development (see Appendix).

We seem to have here a particularly noteworthy instance of a connection that is all too often insufficiently acknowledged and, consequently, insufficiently scrutinized for its implications and applications. While we recognize testing as a goal in and of itself, testing also entails secondary goals, among them 'to achieve a better understanding of how communicative competence is acquired and/or modified by education' (Oller 1986:105). In other words, assessment results, whether covertly or overtly, regularly impact on decisions about who should enter the process of language learning, when and how this should take place, what exactly comprises this activity and, therefore, how its segmentation is to be envisioned and its progression enhanced.

The following discussion will focus on this connection with the aim of addressing some broader issues regarding the testing-curriculum-learner connection. Its immediate impetus is the impact on curriculum of oral proficiency assessment according to the ACTFL mode of assessment in the oral proficiency interview (OPI). But I hope to raise some issues of larger

concern that can help us maximize the call for renewal in the foreign language profession.

Testing and curricular accountability. One way to understand the current healthy tendency to strengthen curricular efforts by stating them in terms of functional ability rather than in terms of seat-time or grammatical components, is by referring to the climate of accountability in American education in general, foreign language instruction in particular. We recognize that, if we wish to bolster our professional respectability and the public's perception of the societal value of our work, we must get used to being judged by the outcomes of that work.

But even without such outside pressures, assessment is valuable for its potential of infusing our discourse with a measure of reality that has frequently been lacking (Omaggio 1987). Pronouncements about what constitutes foreign language study have commonly and for too long been so inappropriate for the task, so vague, or so inflated as to be virtually useless. In that sense, tying insights obtained through proficiency testing to curricular statements has made it possible for many members of the foreign language community to begin to face reality without guilt pangs and to transmit a new sense of the possible and the likely to one another, to the learners, and to administrators at various levels of the educational establishment (Medley 1985). I consider this an eminently valuable turn of events and attribute no small part of the current excitement within the profession to this liberation from false pretenses.

In addition, the concept of assessment in terms of functional use seems finally to have encouraged the profession to tackle one of its perennial problems, namely finding in the *product* of instruction, as ascertained through testing, important clues for organizing and, ultimately, improving the *process* of instruction (Byrnes 1986 and 1987b). The difficulty of such a step should not be understated, but it should not keep us from making differentiated process statements at all. In a sense, all our deliberations, theoretical as well as practical, must contain that loop-back between the outcomes we obtain and the process by which we plan to enhance future learning. Otherwise, engaging in the entire enterprise of testing would take on a rather fatalistic note.

This seems to state the obvious. Yet, one can observe two confounding and contradictory follow-up actions to testing. On the one hand, many instances of testing convey little sense of the appropriateness of, indeed the need to make the connection between product obtained and learning process to be enhanced. Too often the activity of testing is justified solely unto itself. On the other hand, with the advent of the ACTFL Proficiency Guidelines, we now face the opposite scenario, the possible wholesale transfer of a measurement instrument into curricular statements and, perhaps equally dangerously, into materials development. Neither approach is viable. Testing cannot totally disregard the obligation that at least some of its parameters should have the potential of contributing to, even being part of what we understand to be the process of learning (Canale 1984). But the testing cart also cannot and should not be unquestioningly placed in front of the instructional horse.

Interpreting test results. In the current environment, then, where proficiency testing seems to have provided the key impetus for bringing functional language use to the top of the foreign language instructional agenda, the key issue is how to go about *interpreting* testing data so that they can be used to enhance instruction. For any interpretations to be possible, other frames of reference, in addition to those based on outcomes statements, will have to be accessed.

To my way of looking at matters, the first additional frame of reference will be a statement of the theoretical position to which one subscribes vis-à-vis language acquisition. I do not mean to suggest that we currently possess a completely coherent theoretical model, but we must have strong hypotheses regarding numerous issues, among them (1) the influence of L1 on L2 learning, (2) language as closed norm and language as constructed system, (3) learner interlanguage as being norm-deviant or rule-generating, as approximative and constantly being restructured at higher levels or oriented toward mastery of individual chunks where the totality of such mastered chunks would account for mastery of the language, (4) the dominance or subordination of meaning or form during language acquisition.

Almost inescapably, the learner becomes the second frame of reference--the learner as a unique human being with capabilities, beliefs, attitudes, goals, and operating in a given social environment with all its constraints and possibilities. In other words, we must sharpen our awareness of how language learning--and I use the term in the generic sense, not as an opposition to acquisition--takes place and, by extension, how language teaching might best take place under a variety of circumstances for a variety of learners (Schulz 1984).

I suggest that theoretically well-motivated, classroom-based second language acquisition research can provide us with the link between theory and the learner. But the matter is complicated by two factors. First, such research is only now beginning to prosper. Second, already existing evidence does not point in one direction only, precisely because of differing theoretical presuppositions, not to mention different practical research approaches, which inherently govern what kinds of questions can or may be asked, and, consequently, what kind of data can or may be counted as evidence and, finally, what kinds of outcomes assessment are considered appropriate. In this fashion research outcomes are already interpreted outcomes even without interpretive narrative.

In addition, as we turn more and more to qualitative research, including introspection as a tool for capturing important features of learning, the danger is all too great that research designs in themselves are elevated to instructional recommendations, even though they may not at all be viable within a teaching environment.

Yet, we must face these dilemmas. Our task essentially is to analyze in a dispassionate way the available evidence, 'abstracting from individual cases and delicately filtering their implications through the constraints and the opportunities that are operational in the classroom' and, furthermore, respecting with a much greater sense of urgency the 'student population, teaching force, program scope, and, taken in appropriate measure, our desires for the outcomes of foreign language programs' (Byrnes 1987a).

Acknowledging the need for careful weighing of research results is in no way to be equated with the view that the foreign language teaching profession has little empirical evidence for supporting any of its approaches (Lambert 1987). On the contrary, important material has been gathered, but, as often, easily made recommendations are likely to fall significantly short of capturing the complexity of the problem.

The relationship of proficiency testing to curriculum. I indicated earlier that current foreign language instruction can be characterized by an emphasis on functional language ability, captured by the term proficiency in its non-specialized use. However, even its specialized use, referring to the procedure of the oral interview and its rating according to the ACTFL Guidelines, has become an important professional concern. In order to fully understand the opportunities as well as the limitations of applying the ACTFL Proficiency Guidelines and the concomitant testing protocol to curriculum, it is important to analyze their historical background.

Initially, language testing in government circles responded to certain needs which had been broadly circumscribed on the basis of needs assessment. These needs virtually span the entire spectrum of language use, from mere coping in the target language culture to the most sophisticated linguistic, discoursal, cognitive, illocutionary, and sociolinguistic features of language use as they must be manifested in the complex business of representing and furthering the interests of a nation in the diplomatic arena. Consequently, it was only natural that the two end points of the scale were the nonfunctional nonspeaker and the speaker whose versatility and flexibility with the language, for want of a more precise description of this complex phenomenon, was equated with that of the educated native speaker. How a candidate happened to obtain his or her specific configuration of ability between those end points was quite irrelevant for the purposes of assessment. This is the origin of the claim that proficiency testing is independent from the learning or acquisition situation the particular candidate may have experienced.

There are many testing situations where this claim is valid. However, in the strict sense the claim of independence from learning or acquisition history holds only at rather high levels of language ability, and, as mentioned before, even there testing is influenced by the kind of language use for which the government must have evidence. By contrast, when we look more to the middle and lower end of the scale, precisely that range of language abilities which the academic scale, quite rightly, has identified as its primary domain and which has become the cornerstone of curricular work, particularly at the secondary level, the independence of testing procedure from acquisitional history must be and is compromised. Let me explain.

During the administration of an oral proficiency interview, as soon as one must conclude a likelihood that a candidate lacks important elements of command of the target language--and that decision can be made quite rapidly in the vast majority of cases--the further procedure is in essence one of the appropriate selection of topics and contexts and communicative functions which are to be performed. Such appropriateness in topic selection translates into taking recourse to areas that the tester judges to be of interest to the L2

speaker and/or to instructional content presumed to have been part of her or his L2 learning experience.

This constitutes a significant departure from our previous definition of proficiency. In the case of instructed learners, the dependency on prior curricular content is evidenced by testing protocol calling for the inclusion of topical areas such as colors, numbers, days of the week. Such a recommendation occurs not because the ability to handle these topics is inherently more important or more directly indicative of language ability than, say, being able to list kinds of fruits or vegetables, or naming the animals in a zoo or different kinds of kitchen utensils. Instead, it is based on the likelihood that an instructional program covered the former categories relatively early while disregarding, or at least postponing, the others until considerably later. In other words, the less developed the learners' abilities in the L2 the more learners themselves and the presumed conditions and content of their language learning, whether naturalistic or instructed, become an important factor in the testing procedure.

From the standpoint of nonconfrontational testing with a bias for the best, this is as it should be. But we must note a distance to the original concept of proficiency testing, a differentiation which has strong implications for the application of the results of proficiency testing in the lower ranges of the scale --and that kind of testing prevails in the academic setting--for instructional and curricular purposes. Indeed, to separate these essentially different approaches I think it might be much more appropriate to say that, with instructed learners, one engages not so much in proficiency testing in the pure sense but, more likely, in what has been called proachievement testing. Essential qualities of the concept of proficiency and thus proficiency testing are probably not present below the Advanced level. Consequently, recommendations about instructional and curricular implications or applications of what is commonly lumped together as 'proficiency testing' regardless of the speaker's level must consider at least these issues:

(1) A certain circularity of argument between testing evidence, particularly as it pertains to testing content, and curricular goals creeps into testing at the lower end of the scale. Unless there is independent justification for these content areas, curricular guides have a right to and probably an obligation to examine them for their appropriateness in a given instructional setting.

(2) There is a danger of confounding outcomes that are obtained in an essentially curriculum-independent fashion with those that can be obtained through communicatively oriented curriculum-based achievement testing, or proachievement testing, as it has come to be known. Curricular guides which, in contrast with proficiency assessment, must make some reasoned judgment about the length of time necessary for attaining certain goals, could be misled by data from proficiency testing.

(3) One must recognize that, at present, functionally oriented testing is, in a way, gathering evidence on the abilities of learners who, by and large, come from a heavily form-centered instructional environment. Consequently, projections about learner performance in the anticipated functionally oriented curriculum should at least include the potential for upward revision.

(4) Perhaps most importantly, it must be understood that a hierarchy for judging language performance, such as that embedded in the Guidelines, in

no way constitutes a curricular blueprint. Such an application of the Guidelines would be just as unfortunate in its reverse form, namely 'language that learners produce equals language to be taught', as it was in the more customary initial form, namely 'language that was taught equals language students can produce'. Nontransferability of the Guidelines into curriculum would hold even if it should turn out that the assessment hierarchy underlying the Guidelines is not insignificantly related to an acquisitional sequence.

I think it is to the credit of the profession as a whole, that, by and large, the documentation accompanying some of the most prominent curricular initiatives at the state level reflects an understanding on the part of its framers of some of these issues. Those substantive beginnings are all the more important since curricular innovations at the secondary level as they are appearing with increased frequency are the precursors to the next step that is necessary if the promises of insights gained from proficiency-oriented testing are to be fulfilled. I am referring to the yet-to-be accomplished, though clearly overdue, scrutinizing of foreign language curricula at the postsecondary level that is so essential for articulation between secondary and postsecondary programs. And that restructuring, in turn, is indispensable if we are ever to attain the desired upgrading of outcomes of all foreign language instruction that the many commissions and reports challenge us to deliver.

Issues in communicative curriculum development. I suggested that, inherently, proficiency testing in its pure form is applicable within the context of instructed learning in only a small number of testing instances. Consequently, while it is appropriate in instances such as assessing an entire program, secondary or postsecondary, or the assessment of teacher language use as one of many components of certification, its value for curriculum development for the bulk of both secondary and postsecondary language curricula can hardly be in terms of providing direct content information. I emphasize here the wording 'direct content information' to clarify my belief that proficiency testing has only limited applicability to curriculum development. However, I do feel that even an indirect application, namely in terms of an emphasis on functional language use, results in important rethinking of curricular issues (Heilenman and Kaplan 1985). Combined with responsible reflection about appropriate curricular content, it has the potential for making communication not the culminating activity of language learning, to be attempted when the form inventory of a language has been mastered, but its underlying constant at all stages of learning.

The task of curriculum development would thus seem to be twofold: (1) general parameters about the goals of language instruction--in the current professional discussion, functional language use--must be set, along with ways of defining intermediate goals that are indicative of an acceptable rate and scope of progression; (2) specific units of instruction, of various lengths and complexity, must be determined and their sequence motivated by a number of learner-related factors. Among such factors we might find presumed ease of learning under different conditions of learning and with different learners, frequency of use in the L2 social context tied to learner need to know as determined by presumed future L2 use, interest on the part of the learners,

concerns for depth and scope of learning, and the likelihood of specific activities supporting long-term nonreduced and nonstigmatized learning.

As for the first stage, the setting of general parameters, the suggestion offered by Bachman and Savignon (1986) that language abilities should initially be expressed independently of content might be followed. Bachman and Savignon express their hierarchy of language abilities in terms of grammar, cohesion, and sensitivity to register, clearly recognizing the problems inherent in such an approach. Among the difficulties are 'identification and ranking of illocutionary acts in terms of appropriateness and level' and specifying the degree of control and range in terms that are specific enough to distinguish levels clearly and for raters to interpret consistently.

Or one might consider expressing this first stage in terms of the learner's processing capabilities. Roughly speaking, this progression would begin with the word and phrase level, followed by the kernel propositional level reaching up toward the sentence level, and then extend beyond into paragraph and discourse with their respective characteristic cohesive devices. It recommends itself because it is eminently observable in the overall progression learners go through and in their language use at a specific moment along that progression. In particular, it manifests itself in terms of suprasegmental features of intonation and stress, in pause behaviors, and numerous components of fluency, in strategies of meaning transfer, in correction behaviors, in form and length of language chunks, to mention only a few (Byrnes 1987c).

During the second step, after the general features capable of embracing all kinds of curricula have been identified, we must look for specific content units of instruction. These should show increased levels of complexity and respond to a wide variety of social and situational contexts, in terms of the learners as well as in terms of the anticipated environment for L2 language use.

It is encouraging to note that some curriculum proposals which have been offered at the state level (e.g. California, Florida, Indiana, Minnesota, Wisconsin) have accepted the challenge of proficiency-oriented instruction as well as the separate challenge of devising curricular guides that are at once specific enough to fit particular instructional settings and goals but also open enough for modification as teachers gain confidence in their own use of the guides. For example, the Indiana guide arranges instruction around nine broad cultural contexts, such as travel/transportation, school and education, family and home, world of work, and then spirals the curriculum around increasingly complex specific cultural situations in which students might have to function, e.g. getting around in the city. It encourages teachers to "adjust" the competencies to fit varied teaching patterns' and, ultimately, to adapt their evaluation procedures in light of increased skills on their part as well as better teaching materials (4). On a related matter, the University of Minnesota guide for teachers of German states clearly that its efforts at defining a minimal proficiency expected of entering and exiting students in its College of Liberal Arts are to 'be interpreted neither as statements of goals, nor as descriptions of curriculum. Given the gradual, developmental nature of the language acquisition process, the curriculum of a second language program must be targeted well *beyond* these proficiency descriptions. ... Thus, the

guidelines do not in themselves constitute a curriculum but are rather only a framework for a curriculum' (4).

I conclude this section on issues related to the development of communicative curricula with a comment about the dilemma of specificity versus generalizability. As soon as we recognize program- and learner-specific features we are running the risk of losing broad comparability between different programs and approaches. But I believe criteria that express language ability in terms of language-processing capability are broad enough to be applied to different kinds of language development, thus maintaining basic comparability, yet recognizing variation by individual learners, or by setting, or by the goals of language instruction and use (Clahsen 1980).

Summary. In an amazingly short period of time the concept of proficiency, in its generic as well as its specific sense, has captured the interest of the foreign language profession. As is to be expected, the arrival of proficiency in the professional discussion has also led to a range of responses, from unbridled enthusiasm to thorough rejection. To me, our interest in outcomes assessment within the academic setting has resulted in a very healthy awakening of concern for the product of our labors, but also concern for the process by which we arrive at it. Despite many lacunae in the details, a number of key concepts have already been firmly incorporated into our professional discourse, particularly as it relates to classroom methodology and curriculum. Among these I mention calls for personalization, contextualization, curricular reentry and spiraling, interactive and negotiated language use, for a true integration of all skill areas, with listening and reading finally receiving long overdue recognition, for language analysis from the discourse level down, rather than from sentence-level grammar up. I believe that these are concepts that establish valid relationships between product and process. They recognize and respect the complexity of language as they provide a broad frame of reference within which the individuality of our work in each individual classroom with each individual learner can be optimally enhanced.

They also recognize the importance of the learner. Considerations about program breadth and scope, methodological preferences, curricular sequencing, materials selection, about balancing between different skill areas, or selecting appropriate goals due to maturational considerations, ultimately always lead back to individual learners with individual learning styles. Of course, we know far too little about these issues, and we will likely be in that position for some time to come since the possible configurations are essentially infinite. However, there are some major shared parameters, which the study of psycholinguistic processing, particularly in classroom-centered second-language acquisition research, might help to identify more closely. In the end the classroom with its learners is the locus of decision making, and we as teachers are the ones who are challenged and empowered to make these decisions.

Appendix: ACTFL Proficiency Guidelines

The 1986 proficiency guidelines represent a hierarchy of global characterizations of integrated performance in speaking, listening, reading and writing. Each description is a

representative, not an exhaustive, sample of a particular range of ability, and each level subsumes all previous levels, moving from simple to complex in an 'all-before-and-more' fashion.

Because these guidelines identify stages of proficiency, as opposed to achievement, they are not intended to measure what an individual has achieved through specific classroom instruction but rather to allow assessment of what an individual can and cannot do, regardless of where, when, or how the language has been learned or acquired; thus, the words "learned" and "acquired" are used in the broadest sense. These guidelines are not based on a particular linguistic theory or pedagogical method, since the guidelines are proficiency-based, as opposed to achievement-based, and are intended to be used for global assessment.

The 1986 guidelines should not be considered the definitive version, since the construction and utilization of language proficiency guidelines is a dynamic, interactive process. The academic sector, like the government sector, will continue to refine and update the criteria periodically to reflect the needs of the users and the advances of the profession. In this vein, ACTFL owes a continuing debt to the creators of the 1982 provisional proficiency guidelines and, of course, to the members of the Interagency Language Roundtable Testing Committee, the creators of the government's Language Skill Level Descriptions.

ACTFL would like to thank the following individuals for their contributions on this current guidelines project: Heidi Byrnes, James Child, Nina Levinson, Pardee Lowe, Jr., Seiichi Makino, Irene Thompson, and A. Ronald Walton.

These proficiency guidelines are the product of grants from the U.S. Department of Education.

Generic Descriptions: Speaking.

Novice. The Novice level is characterized by the ability to communicate minimally with learned material.

Novice-low. Oral production consists of isolated words and perhaps a few high-frequency phrases. Essentially no functional communicative ability.

Novice-mid. Oral production continues to consist of isolated words and learned phrases within very predictable areas of need, although quality is increased. Vocabulary is sufficient only for handling simple, elementary needs and expressing basic courtesies. Utterances rarely consist of more than two or three words and show frequent long pauses and repetition of interlocutor's words. Speaker may have some difficulty producing even the simplest utterances. Some Novice-Mid speakers will be understood only with great difficulty.

Novice-high. Able to satisfy partially the requirements of basic communicative exchanges by relying heavily on learned utterances but occasionally expanding these through simple recombinations of their elements. Can ask questions or make statements involving learned material. Shows signs of spontaneity although this falls short of real autonomy of expression. Speech continues to consist of learned utterances rather than of personalized, situationally adapted ones. Vocabulary centers on areas such as basic objects, places, and most common kinship terms. Pronunciation may still be strongly influenced by first language. Errors are frequent and, in spite of repetition, some Novice-High speakers will have difficulty being understood even by sympathetic interlocutors.

Intermediate. The Intermediate level is characterized by the speaker's ability to: (1) create with the language by combining and recombining learned elements, though primarily in a reactive mode; (2) initiate, minimally sustain, and close in a simple way basic communicative tasks; and (3) ask and answer questions.

Intermediate-low. Able to handle successfully a limited number of inter-active, task-oriented and social situations. Can ask and answer questions, initiate and respond to simple statements and maintain face-to-face conversation, although in a highly restricted manner and with much linguistic inaccuracy. Within these limitations, can perform such tasks as introducing self, ordering a meal, asking directions, and making purchases. Vocabulary is adequate to express only the most elementary needs. Strong interference from native language may occur. Misunderstandings frequently arise, but with repetition, the Intermediate-Low speaker can generally be understood by sympathetic interlocutors.

Intermediate-mid. Able to handle successfully a variety of Mid un-complicated, basic and communicative tasks and social situations. Can talk simply about self and family members. Can ask and answer questions and participate in simple conversations on topics beyond the most immediate needs; e.g., personal history and leisure time activities. Utterance length increases slightly, but speech may continue to be characterized by frequent long pauses, since the smooth

incorporation of even basic conversational strategies is often hindered as the speaker struggles to create appropriate language forms. Pronunciation may continue to be strongly influenced by first language and fluency may still be strained. Although misunderstandings still arise, the Speaker can generally be understood by sympathetic interlocutors.

Intermediate-high. Able to handle successfully most uncomplicated high communicative tasks and social situations. Can initiate, sustain, and close a general conversation with a number of strategies appropriate to a range of circumstances and topics, but errors are evident. Limited vocabulary still necessitates hesitation and may bring about slightly unexpected circumlocution. There is emerging evidence of connected discourse, particularly for simple narration and/or description. The Intermediate-High speaker can generally be understood even by interlocutors not accustomed to dealing with speakers at this level, but repetition may still be required.

Advanced. The Advanced level is characterized by the speaker's ability to: (1) converse in a clearly participatory fashion; (2) initiate, sustain, and bring to closure a wide variety of communicative tasks, including those that require an increased ability to convey meaning with diverse language strategies due to a complication or an unforeseen turn of events, and (3) satisfy the requirements of school and work situations; and narrate and describe with paragraph-length connected discourse.

Advanced. Able to satisfy the requirements of everyday situations and routine school and work requirements. Can handle with confidence but not with facility complicated tasks and social situations, such as elaborating, complaining, and apologizing. Can narrate and describe with some details, linking sentences together smoothly. Can communicate facts and talk casually about topics of current public and personal interest, using general vocabulary. Shortcomings can often be smoothed over by communicative strategies, such as pause fillers, stalling devices, and different rates of speech. Circumlocution which arises from vocabulary or syntactic limitations very often is quite successful, though some groping for words may still be evident. The Advanced-level speaker can be understood without difficulty by native interlocutors.

Advanced-plus. Able to satisfy the requirements of a broad variety of everyday, school, and work situations. Can discuss concrete topics relating to particular interests and special fields of competence. There is emerging evidence of ability to support opinions, explain in detail, and hypothesize. The Advanced-Plus speaker often shows a well developed ability to compensate for an imperfect grasp of some forms with confident use of communicative strategies, such as paraphrasing and circumlocution. Differentiated vocabulary and intonation are effectively used to communicate fine shades of meaning. The Advanced-Plus speaker often shows remarkable fluency and ease of speech but under the demands of Superior-level, complex tasks, language may break down or prove inadequate.

Superior. The Superior level is characterized by the speaker's ability to: (1) participate effectively in most formal and informal conversations on practical, social, professional, and abstract topics; and (2) support opinions and hypothesize using native-like discourse strategies.

Superior. Able to speak the language with sufficient accuracy to participate effectively in most formal and informal conversations on practical, social, professional, and abstract topics. Can discuss special fields of competence and interest with ease. Can support opinions and hypothesize, but may not be able to tailor language to audience or discuss in depth highly abstract or unfamiliar topics. Usually the Superior level speaker is only partially familiar with regional or other dialectical variants. The Superior level speaker commands a wide variety of interactive strategies and shows good awareness of discourse strategies. The latter involves the ability to distinguish main ideas from supporting information through syntactic, lexical and suprasegmental features (pitch, stress, intonation). Sporadic errors may occur, particularly in low-frequency structures and some complex high-frequency structures more common to formal writing, but no patterns of error are evident. Errors do not disturb the native speaker or interfere with communication.

References

- Bachman, Lyle F., and Sandra J. Savignon. 1986. The evaluation of communicative language proficiency: A critique of the ACTFL oral interview. *Modern Language Journal* 70.4:380-90.
- Byrnes, Heidi. 1987a. Second-language acquisition research: Does it suggest program goals? *ADFL Bulletin* 18.3:35-38.

- Byrnes, Heidi. 1987b. Proficiency as a framework for research in second language acquisition. *MLJ* 71.1:44-49.
- Byrnes, Heidi. 1987c. Speech as process. *Foreign Language Annals* 20.4:301-10.
- Byrnes, Heidi. 1986. Second-language acquisition: Insights from a proficiency orientation. In: *Defining and developing proficiency: Guidelines, implementations and concepts*, ed. Heidi Byrnes and Michael Canale. Lincolnwood, Ill.: National Textbook Company. 107-31.
- Canale, Michael. 1984. Considerations in the testing of reading and listening proficiency. *Foreign Language Annals* 17.4:349-57.
- Clahsen, Harald. 1980. Psycholinguistic aspects of L2 acquisition. In: *Second language development: Trends and issues*, ed. Sascha Felix. Tübingen: Gunter Narr. 57-79.
- Galloway, Vicki. 1987. From defining to developing proficiency: A look at the decisions. In: *Defining and developing proficiency: Guidelines, implementations, and concepts*, ed. Heidi Byrnes and Michael Canale. Lincolnwood, Ill.: National Textbook Company. 25-73.
- Heilenman, Laura K., and Isabelle M. Kaplan. 1985. Proficiency in practice: The foreign language curriculum. In: *Foreign language proficiency in the classroom and beyond*, ed. Charles J. James. Lincolnwood, Ill.: National Textbook Company. 55-78.
- Indiana curriculum guide. 1986. A guide to proficiency-based instruction in modern foreign languages for Indiana schools. Indiana Department of Education.
- Lambert, Richard D. 1987. The case for a National Foreign Language Center: An Editorial. *Modern Language Journal* 71.1:1-11.
- Medley, Frank W., Jr. 1985. Designing the proficiency-based curriculum. In: *Proficiency, curriculum, articulation: The ties that bind*, ed. Alice C. Omaggio. Middlebury, Vt.: The Northeast Conference. 13-40.
- Minnesota curriculum guide for German. 1986. Foreign language proficiency standards for entrance and graduation. College of Liberal Arts. A guide for teachers of German. University of Minnesota.
- Oller, John W., Jr. 1986. Communication theory and testing: What and how. In: *Toward communicative competence testing: Proceedings of the Second TOEFL Invitational Conference*, ed. Charles W. Stansfield. Princeton, N.J.: Educational Testing Service. 104-55.
- Omaggio, Alice C. 1987. Teaching language in context. *Proficiency-oriented instruction*. Boston: Heinle and Heinle.
- Schulz, Renate A. 1984. Language acquisition and syllabus design: The need for a broad perspective. *ADFL Bulletin* 15.3: 1-7.

Assessment, articulation, accountability: New roles for the language lab

Sue K. Otto

The University of Iowa

For years, deans and department heads have been grappling with cost-effective ways to improve foreign language education. Inspired by recent faculty interest in communicative methodologies and the promise of new instructional technologies to help them toward their goals, administrators have decided to replace their outmoded language lab facilities with modern audio equipment and to diversify their operations to include computer, video, and interactive audio and video technologies.

Following the traditional lead of textbook publishers, too many teachers still regard language laboratory materials--audio, video, and computer software--as ancillary to the 'real' courseware (the book and accompanying teacher strategies and techniques). Certain preconceptions remain entrenched: audio (yawn), video (Friday's movie treat), computer (still not convinced), interactive video and audio (very expensive). Although a vast body of literature about proficiency-based and communicative methods exists, relatively little consideration has been given to detailing how laboratory technology can be integrated into the curriculum most effectively to enhance teaching and learning. Yet if we look at several currently important language issues--assessment, articulation, and accountability--and if we consider the strengths of present technologies, we can cast the language lab in a number of convincing new roles.

In this article I focus on projects at the University of Iowa to illustrate new directions in language learning technologies. Many colleagues at other institutions are engaged in similar projects, working toward similar goals. My Iowa bias results from familiarity with our own programs and not from contempt for others.

Assessment and technology. In some departmental programs testing fossilized during the grammar-translation era. The placement tests for foreign languages we administer to incoming freshmen at the University of Iowa are a case in point. Although the course exams from most of the departments cover grammar, vocabulary, reading, and listening comprehension and employ a good variety of objective item formats, the placement exams are traditional discrete-point, multiple choice grammar/vocabulary tests. Their only saving grace is that they are machine-scorable and the Exam Service administers them for the departments. They are no longer reliable instruments for placement, since our goals for language study have changed. Students who are beyond the first year are expected to have achieved a certain level of skill in

speaking, listening, reading, and writing in order to function successfully; knowledge of grammar is not a reliable yardstick for measuring these skills.

Traditional paper-and-pencil testing methods have broadened to test more than grammar and vocabulary facts as departments turn to methods that emphasize both productive and receptive skills. While most of Iowa's course exams and many of the standardized tests on the market still have a significant portion of traditional grammar and vocabulary questions, they have broadened to include reading and listening questions (some cued with pictures and recordings), free-response writing exercises, and, occasionally, a speaking component. Among others the Advanced Placement (AP) foreign language exams, the Connecticut state foreign language exams developed by ETS, and the University of Minnesota's Entrance Standard and Graduation Requirements tests exemplify a new phase of improved instruments for comprehensive language testing. Yet in most programs a serious gap between teaching and testing remains for the productive skills, particularly speaking. We have been circling in a holding pattern for too many years: the 'easy' areas of grammar, vocabulary, and reading comprehension receive the most thorough testing; listening comprehension and writing are tested in a more cursory fashion; and speaking--the most anxiety-producing mode for students and the most difficult to assess for teachers--is often left out altogether. Language books and journals are filled with 'blue sky' rhetoric from the communicative and proficiency experts about improvements in teaching methods. Why haven't we made more progress in testing all the important aspects of language?

Lange (1987) candidly discusses the prevalent characteristics of language learning in universities that impede the implementation of proficiency-oriented teaching. The teaching staff is composed largely of graduate students who have no formal training in language teaching (and sometimes no interest in it). Program coordinators, who shape and maintain the programs, often come to the job with a preparation in literature rather than in pedagogy and are not rewarded for research and development in pedagogy. Colleagues in literature often disdain language teaching as beneath their dignity. Departments lack a unifying curriculum design that gives direction to the program as a whole. In spite of all the lip service to communicative competence and proficiency, most of the learning materials and tests at both high school and university levels are still grammar/translation-oriented. Classes are too large for teachers to effectively apply techniques.

The Oral Proficiency Interview (OPI) presents the perfect example of an ideal defeated by reality. While many are convinced that the OPI is the best solution we have for measuring speaking abilities, it will never be a widely used technique in high school or university language programs. Assessment of student speaking ability is extremely time-consuming, requiring many teacher hours to obtain valid samples and to grade them. Educating the teaching staff in complex presentation and evaluation techniques of the OPI is a challenging, often futile task; and invariably consistency of evaluation comes into question. It is not easy to administer and grade speaking tests, even if there are relatively few students to be examined; in departments with hundreds or thousands of students the overwhelming burden of testing them one-on-one dampens the enthusiasm of even the most ardent advocates of the

OPI. Is there no answer to the charge that the OPI, while effective, is too resource-intensive to be practical?

In spite of all these negative arguments, leaders of the language profession continue to encourage methods that produce students who can function in the language; and they still push for requirements based on competence rather than seat time. Some universities have already pioneered testing in the hazardous terrain of the proficiency-oriented curriculum, notably the University of Minnesota (Arendt et al. 1988; Lange 1987), the University of Tennessee (Rogers et al. 1988), the University of Pennsylvania (Schulz 1988), and the University of South Carolina (Medley 1985). Now more institutions have joined these universities in modernizing the goals and administration of language programs to reflect proficiency-related priorities. At the University of Iowa we have embarked on a two-year Foreign Language Assessment Project to explore new testing methods, recommend placement policy, and devise new placement exams. The Project, codirected by Irene Wherritt (Department of Spanish and Portuguese) and T. Anne Cleary (Director of the University of Iowa Evaluation and Examination Service), is the joint venture of representatives from the foreign language departments, personnel from the University of Iowa Examination and Evaluation Service (which is hiring a new staff member expressly for foreign language testing), and the Language Media Center. The product of our labor will bear the stamp of our collective expertise and that will mean a greater role for the language laboratory in language assessment.

As Larson (1986) notes, interest in shifting language testing to more appropriate technologies than paper and pencil has already produced a number of promising examples of our immediate future in technology-based interactive testing. Imaginative use of technology has the power to improve both the practical aspects of testing for the test-givers (especially in terms of administrative efficiency, security, and consistency with new proficiency-oriented methods) and the testing experience for the examinee. Microcomputers form the basis for the most convincing new learning and testing strategies to be introduced in the language lab in recent years: computerized adaptive testing, interactive reading, and interactive listening comprehension with videodisc.

Computer adaptive tests provide an effective and efficient way to test grammar, vocabulary, and reading skills. At Brigham Young University Larson (1989) developed a program called S-CAPE: Spanish Computerized Adaptive Placement Exam. The program functions with a large database of multiple choice vocabulary, grammar, and reading items that have been analyzed, calibrated, and coded for type and difficulty level. As a student takes the exam, the computer adapts to that student's level of ability through a series of level checks and probes like those typical of the OPI. In terms of reliability and validity the exam fared well in studies comparing it to national standard paper-and-pencil tests. There are a number of important benefits to the approach. Testing time is greatly reduced--averaging twenty minutes for the exam, even though students are allowed to work at their own pace with no time constraints. Because of the tailored, streamlined process students are less bored and frustrated by many questions that are too hard or too easy. The results are immediately available to the student at the end of the test.

The administrative burden is on the computer (giving instructions, scoring and reporting the results) so fewer test administrators are necessary. Test security is enhanced since each test is unique. Given the computer resources that are currently available in language labs on many campuses, this kind of testing could be immediately implemented with very little additional investment in hardware.

One of the most scathing indictments of computer software has been that much of it is just electronic page turning. However, projects like the Interactive Reading Project (IRP), directed by Geoffrey Hope (Department and French and Italian) and George S. Hirst (CONDUIT) at the University of Iowa, are exploring ways to strengthen and assess reading skills through a variety of interactions with authentic texts. By providing work with French, German and Spanish texts on the Macintosh, the interactive reading environment offers the reader a setting for both extensive and intensive reading supported by a variety of help features and activities that keep the student constantly interacting with the text meaningfully. On the computer the text becomes malleable and responsive to student input during the activities. Temporary highlighting of words or phrases can be used to draw attention to a thread of meaning and related concepts or to provide a means of feedback for wrong answers. The mouse can be used to point at parts of the text for responding to questions or for dragging words from the text to another part of the screen or vice-versa. Graphic and written commentaries, glosses, and other kinds of help on specific text elements are seen only if they are needed by the reader. Students may follow a course prescribed by their instructor or explore freely, switching between texts and activities. Since it is computer-based, keeping records on student use and performance can be accomplished with ease. From this we begin to envision a near-future reading test built around a large database of texts, perhaps on CD ROM, and a suite of skill-specific programmed activities that employ the computer-adaptive interaction model, all tracked, scored, and reported by the computer. We have the technology; the challenge is to deliver software that exploits it appropriately.

Computer-controlled media--especially videodisc--further expands the testing horizon for listening comprehension. One of the goals of the University of Iowa's Project for International Communication Studies (PICS), funded by grants from the Department of Education and the Annenberg/CPB Project, is to develop foreign language videodiscs and computer software for authentic television materials. The programs produced to date include The Listening Tool (a playback-control program), Plot Puzzle (a game to unscramble clips from a video), Cloze Captions (cloze encounters with video), and Comprehension Check (question/answer format cued with video). All the tools and tasks are designed to move students away from what might be termed the 'couch potato' mode of listening toward intensive interaction motivated by the video material, targeting both intensive and global objectives. In a field previously dominated by audio and multiple choice paper-and-pencil exercises, we are now giving students richer, more realistic listening experiences and a much broader range of activities for strengthening their listening abilities.

Building on the original video, PICS is experimenting with a number of strategies for exploiting the power of this medium, including production of alternate audio tracks (simplified as well as more carefully enunciated versions of the original), placing transcript and keyword frames on the videodisc, using subtitles in the original language, and overlaying other computer graphics on the video to label and highlight parts of the image. Just as with the IRP materials, techniques for learning concepts and skills under computer control can be used effectively for testing as well. By combining the efforts of IRP and PICS, we hope to discover the extent to which the receptive skills are interrelated and whether they can be mutually beneficial in building and testing proficiency.

It is important to maintain consistency between the use of technology in teaching and learning activities and its use in testing. If video is used for in-class teaching and for homework assignments, it should be used for testing as well--and vice-versa. The same holds true for audio, computer, and any other medium we depend on in our teaching for specific skill-building activities. Our campus testing specialists have stressed that testing should not be separate from instruction; the more often teachers avoid the traditional anxiety-ridden hour of class set aside just for 'the test,' the better. Constant assessment of performance on assignments can replace much of the distinct testing phase in a course. When computers are involved in the learning process, evaluation of skill levels and reporting can often be an integral part of the software.

Testing the productive skills of speaking and writing is a tough challenge for humans and an even tougher challenge for machines. Certainly the field of computer artificial intelligence (AI) and natural language processing is not yet up to the rigors of assessment. Though it is unlikely that computers or computer-controlled media will make any real progress in relieving us of the burden of actually assessing the productive skills, even now we have indications that technology can make a significant contribution in administering the tests.

Audio tape, the most traditional and humble medium in the language lab, can eliminate a number of the drawbacks of oral testing. Testing oral proficiency by the best instrument we currently have--the Oral Proficiency Interview--demands a talented interviewer and a great deal of time both to perform the interview and to rate the resulting sample. Using this technique to test massive numbers of students proves to be too much of a drain on most university programs. Are we therefore doomed to slight assessment of oral proficiency forever, even though we claim that speaking competence is essential? Until AI delivers a better solution, some compromise is indicated: we must streamline the process of getting the sample and resign ourselves to providing human hours to handle the assessment. A carefully designed audio-taped interview reduces the one-on-one contact required to interview each examinee personally and results in standardized content that is perhaps easier to judge. Working examples of this are now in use.

On our campus, ESL is the only language program that dedicates the necessary personnel resources to test speaking proficiency. Each semester, the Exam Service administers in our language laboratory the SPEAK test, which was produced by ETS as a measure of oral proficiency in English. The ESL

teachers, who have been trained in assessment techniques for this exam, then rate the samples recorded in the lab; this is a time-consuming and highly complex task, but a critical one. All foreign teaching assistants are required to take this exam and score at a level prescribed by the kind of teaching duties they are to perform; those who fail to demonstrate the prescribed proficiency level, aren't allowed to teach. Although there is some grumbling from both sides about the content and conditions of taking the test, the approach works well enough to justify administering it, since there is an immediate practical need for assessment of the ability to communicate orally.

The Portuguese Speaking Test (PST) (Stansfield et al. 1988), an audio cassette-based simulation of the OPI developed by the Center for Applied Linguistics, has been successful in replacing the person-to-person part of the interview. Described in the test manual as a 'semi-direct test of oral proficiency,' the test adheres to ACTFL/ILR standards for intermediate and advanced levels in probing for linguistic and sociolinguistic evidence of proficiency; and it is scored using the ACTFL/ILR Guidelines. According to studies comparing the validity and reliability of the PST and the OPI, there was a very high correlation between the results of the two tests. Stansfield (1989) observes that the recorded interview holds some advantages over the live interview. In addition to saving personnel time to conduct the interview, the quality of the interview is no longer dependent on such personal variables as how talented the interviewer is at eliciting a ratable sample, or how she feels that day. Moreover, the taped interview often elicits a longer sample to rate; live interviewers tend to intervene more frequently in the course of the interview and this results in less interviewee speech for the time spent interviewing. Modern lab equipment facilitates recording only the examinee's responses on the tape, thus avoiding the extra time it would take a rater to scan past the examiner's cues. Only the best raters are assigned to assess the samples once the tapes come back to CAL. This kind of personnel specialization may be in store for universities if we decide to pursue testing of oral abilities for multisection foreign language courses in the future.

Everyone knows the mechanical benefits of word processing and its accompanying resources, including spelling checkers, thesauri, and mechanisms for automatically generating a table of contents or an index. Some software attempts to check the style and syntax of the prose as well. However, as in the case of assessing speech, the machines can help in getting a good sample, but the assessment of the writing itself has to be done by humans. The advances in computer understanding of natural language have been largely disappointing, as Underwood's survey (1987) of AI and computer-assisted language learning shows. Glimmers of hope spring from projects like Syncheck, Ruth and Alton Sanders' German composition processor that parses elementary German; but defining the role that language lab computers and other media might play in the assessment of language production is clearly a task for the future.

Personally, I am a firm skeptic about AI. Stewart Brand (1987) has reinforced my doubts in *The Media Lab: Inventing the Future at M.I.T.* Brand quotes Marvin Minsky, one of AI's founding fathers: 'Anything you hear about computers or AI should be ignored because we're in the Dark Ages. We're in the thousand years between no technology and all technology. You

can read what your contemporaries think, but you should remember they are ignorant savages' (p. 104). Negative feelings about AI aside, reading *The Media Lab* has reinforced some of the some of my own more positive observations about trends in learning laboratory technology. Ten years ago the founders of the M.I.T. Media Laboratory saw in the merging of the major media industries--broadcasting and film, print and publishing, and computing--a point of departure for much more powerful media in the future. As in the real world, the traditionally distinct forms of laboratory technology--print, audio, video, and computer--have begun to blend together under computer control. Some of the newest audio labs are computers in disguise that broadcast and capture speech in digital form, perform random-access computing functions, and access information stored in both magnetic and optical (videodisc and CD-ROM) form. Instead of maintaining separate facilities that address specific tasks and skills--an audio lab, a computer lab, a video lab, and a library--the lab will be a synthesis of all of these, and assessment will be one of its primary missions.

Articulation and the lab. Having examined the assessment potential of lab technology, let us now turn to articulation and the issue of whether the lab has a role in solving the problems caused by the lack of articulation between high school and college language instruction. For years, the problem of articulation has plagued language students and advisors (Lange 1988). Beginning university language classes are populated by students who have already studied the language for several years but who are too insecure or too lazy to enter a higher-level language course. The result is boredom for the 'false beginner,' intimidation for the real beginners in the class, and overcrowding in lower-level classes. As universities establish entrance requirements for foreign languages in an effort to force students to study beginning language in high school and focus university language study at a higher plane, we must find practical solutions to the problems of articulation and placement. Whether the entrance requirement is measured by a minimum level of proficiency or in years of seat time, colleges must accurately assess incoming students and place them appropriately for their functional abilities in the language.

Starting in 1990, one of the entrance requirements for the University of Iowa will be two years of high school foreign language. Once at the university, most undergraduate students are required to complete two years of foreign language study (through the fourth semester) to graduate. One of the goals of the Foreign Language Assessment Project at Iowa is to produce proficiency-based placement tests and to establish both the 'sticks' to dissuade students from starting over and the 'carrots' that will encourage them to pursue higher-level language courses. For any system like this to function properly, we must work on articulation between the high school and college levels.

Of course, many of the problems of articulation would be solved if the profession could set standards at a national level that would prescribe common methods for teaching and testing foreign language skills. The goal of proficiency in all four skill areas--speaking, listening, reading, and writing--has been adopted by many teachers. However, as Valdman (1988) observes, we

have barely begun work on the elaboration of a national proficiency-based standard. In fact, because there is so much controversy over proficiency-oriented teaching and testing methods, it seems unlikely that we will ever achieve such a standard. As assessment instruments improve and instructional technologies evolve, however, we will be able to cope with articulation in more creative and economical ways. We will increase not only the amount of testing we do in the lab, but how much independent remediation takes place there as well.

Placement exams and tests for credit are designed to report in a rather generalized way how well prepared or how proficient a student is in the language. Currently, little use is made of such exams for diagnostic purposes. A decade ago a faculty member at the University of Iowa devised a computerized diagnostic grammar test that produced for the student a kind of remedial guide--a list of grammar points to be reviewed and the locations in the first year textbook where explanations of the grammar could be read. Though the test did not survive, it is time to resurrect its basic premise of using a test to prescribe a course of remediation. To use a more recent example, with some modifications the computerized adaptive test could analyze a student's strengths and weaknesses and indicate what areas need remedial work. Based on the test outcome, the student might be directed to practice more on specific grammar points, to build vocabulary, to strengthen specific listening comprehension or reading skills, to work on pronunciation, or to improve speaking skills. In all but the last case, in which only a human interlocutor will do, the student is then directed to the language laboratory where appropriate audio, video, and computer technologies provide the means to remedy specific deficiencies at the student's own pace. The responsibility for remedial work belongs to the student, not some unfortunate first-year foreign language TA.

Another approach for bringing students 'up to speed' was designed five years ago in a proposal to institute an 'Express Language Track' for the University of Iowa. With this plan students would be encouraged to build on past work rather than discarding it. After a careful assessment of skills, a student might be placed in a higher level course and enroll concurrently in a two-hour 'Express Track' course of independent study that targeted individual weaknesses. Although an instructor monitored and advised students along the way, a multimedia computer workstation (including interactive video and audio) would deliver the actual instruction. Five years ago this was an idea whose time had not come, if granting agencies are any indication. Yet the proposed solutions for teaching to individual deficiencies are even more valid today when we consider the improved potential of the technology to offer a variety of language learning experiences and to assess learning progress.

One of the most obvious strategies to combat the pain of articulation is to improve communication between high schools and universities. The formulation of general standards and goals should at least be attempted at the state-wide level. High schools, colleges, and universities must collaborate to agree on their expectations and goals and outline them in concrete terms. While we work toward this ideal, there are other less sweeping cooperative measures we can take.

Once universities have established entrance and exit requirements and have tests devised for enforcement of these requirements, we should share them with the high schools so that there are fewer surprises later. What better way to illustrate how our advanced programs work? High school teachers should be able to experience the testing procedures firsthand during workshops or visits to the language lab. Providing direct experience to students would be feasible only in the high school setting, which may not have the same technological resources that the universities do, especially those that are administered interactively. Exam samples could be 'captured' from interactive workstations using more standard media, such as audio or video tapes that could be distributed to high schools; this service is the logical extension of the language laboratory's role in the assessment process.

Accountability: The language laboratory perspective. Moving beyond assessment and articulation, we can identify other trends that have redefined the role of the language laboratory. It is time to finally come to terms with the cliché of the 'failure of the language lab' and move ahead with our best efforts to improve instruction through technology. Language media professionals must always be conscious of the fact that, like classroom teachers, they should be held accountable for their methods and goals. Can we justify what we are doing and how we are doing it? We must constantly scrutinize and reevaluate what we are doing, expecting that lab services and functions will change as teaching methods and technology evolve.

One of the most striking changes in the lab in recent years has been in laboratory directorship, which has been transformed from caretaker/technician to full-fledged language professional. Over the last several years, I have seen only one advertisement for a lab director position that listed the ability to do equipment repairs as a qualification for the job. There also appears to be a decreasing reliance on drafting a junior faculty member as an approach to filling the directorship of the lab. Universities now hope to ensure the success of the expensive new labs by hiring dynamic language media specialists who can collaborate on an equal footing with faculty to build sound technology-based programs. Recent job announcements for lab director positions reflect much higher expectations for applicants than in the past. Among the commonly listed credentials are: (1) a higher degree--Ph.D. or M.A., (2) experience in foreign language teaching and knowledge of the new methodologies, (3) expertise in new technologies and their applications to foreign language learning, (4) vision and leadership abilities, including skills in working with faculty and securing grants, (5) demonstrated ability to administer a service-oriented unit. In defining the role of the lab, research and development of technology-based language learning programs now have a high priority.

The look and feel of the lab itself have also changed. The teacher-directed lab hours formerly spent by classes in monitored oral drills have become a thing of the past. Lab facilities are now designed to accommodate and encourage more independent learner-directed interaction with the technology. As we build our information base through access to mass magnetic and optical storage and to worldwide communication links, we build

our labs into a richer and more appealing environment for both structured and exploratory study modes.

Physical facilities are wasted unless we foster the thoughtful integration of appropriate technology in the curriculum. Technology-based activities directed by the teacher in the classroom and by the learner working independently in the lab must be carefully prepared to fit into teaching as well as testing. Our perspectives on technology have changed. We no longer look to one technology (audio) to serve all of language learning, and we have learned not to expect miracles from new hardware for which techniques and software have not yet matured. Each of the technologies that remain in the lab are there because of their recognized strengths in language learning.

Once the old prejudices engendered by past misuses of audio technology are erased from our minds, the audio lab reveals itself to be a refined, useful tool that can be used to motivate students through creative, lively activities like those outlined by Ely (1984), Johnson and Dvorscak (1984), and Stone (1988). The exercises produced by all these authors were developed to conform to their own methods and curriculum content and to strengthen specific subskills in speaking and listening comprehension. A far cry from blind implementation of publishers' tape programs, such tailoring of tasks and content is critical to the use of this technology.

For a time, critics warned us that computer-assisted language instruction was destined to repeat the history ('failure') of the audio lab--another case of an expensive new technology in search of applications. We have now begun to discover the best strategies for using the computer to our best advantage. Although the use of the computer as an electronic workbook has been severely maligned by those who think we ought to be conversing with computers, at the University of Iowa paper workbooks have effectively been replaced by much more effective computer programs administered by the Language Media Center. The programs diagnose students' errors helpfully and can report the student's usage record to a specific instructor. The students get practice that prepares them for tests as well as immediate performance feedback during their session; and the instructors are grateful to be rescued from the tedious chore of grading written exercises. During the fall 1988 semester, University of Iowa language students accessed the language CAI programs 28,913 times, spending 15,460 hours doing grammar and vocabulary exercises in Spanish, French, German, and Italian. These numbers are even more astonishing when one discovers that only about half of the students were required to spend time on CAI. The success and acceptance of computerized drill and practice have encouraged us to explore other ways to exploit the power of computing in more proficiency-oriented activities.

Similar creative exploration has been essential to the effective implementation of video for foreign language instruction. For years, movies and videotapes were the junk-food of the curricular diet. It was easier to teach with movies: students were delighted to be treated to a day off from the real 'meat' of the course and the teacher didn't have to cook. A perusal of the excellent practical guide to using video in the classroom, 'The Video Connection: Integrating Video into Language Teaching' (Altman 1989), is convincing evidence that video has matured into a genuinely effective medium for language instruction both in class and in the language laboratory.

Interactive videodisc--combining all the benefits of computer and videodisc technologies--is the innovation that has attracted the most attention these days. Interactive video has enormous potential to improve individualized language instruction in the language laboratory. Early projects in interactive video (Gale 1989; Rivera-LaScala 1989; Verano 1989; Rubin 1989) have been so compelling that some institutions have been prompted to consider discarding all their other lab technologies and converting to interactive videodisc lab workstations. At this point the decision to completely abandon older technologies in favor of videodisc is a mistake because the videodisc medium is still an underdeveloped technology with very little available software. It is not yet clear, in fact, that videodisc will withstand the strong competition of other new technologies. Until we reach a stabilized total technology sometime in the next millenium, we must exploit all of the technologies we have in the lab in ways best suited to their current state.

Conclusion. The evolution of the new-age language laboratory has been shaped largely by the demands of new methods in language pedagogy finding resonance in new and compelling technical tools: modern audio and video equipment, mass magnetic and optical storage media, and computers. Through these splendid technologies, the lab is delivering stimulating and challenging media-based materials that respond to the needs of departmental curricula as well as to those of the independent learner. Staffed by professionals, the lab offers--in addition to technical expertise--pedagogical consultation, research and development support, and training for faculty and student users. If its resources are used wisely, the language lab will play valuable roles in proficiency oriented assessment, articulation, research and development, and enhancement of teacher and student productivity.

References

- Altman, Rick. 1989. *The video connection: Integrating video into language teaching*. Boston: Houghton Mifflin.
- Arendt, Jermaine D., Dale L. Lange, and Ray Wakefield. 1986. Strengthening the language requirement at the University of Minnesota: An Initial Report. *Foreign Language Annals* 19.2:149-56.
- Brand, Stewart. 1987. *The media lab: inventing the future at M.I.T.* New York: Penguin Books.
- Ely, Philip. 1984. *Bring the lab back to life*. New York: Pergamon Press.
- Gale, Larrie E. 1989. Macario, Montevidisco, and interactive Digame: Developing interactive video for language instruction. In: *Modern technology in foreign language education: Applications and projects*. Wm. Flint Smith, ed. Lincolnwood, Ill.: National Textbook Company. 235-248.
- Johnson, Lathrop P., and Andrea Dvorscak. 1984. Is there a doctor in the language lab? or: First aid for lab programs. *Unterrichtspraxis* 17.1:28-32.
- Lange, Dale L. 1987. Developing and implementing proficiency oriented tests for a new language requirement at the University of Minnesota: Issues and problems for implementing the ACTFL/ETS/IRL Proficiency Guidelines. *Proceedings from the Symposium on the Evaluation of Foreign Language Proficiency*. Bloomington, IN: Indiana University.
- Lange, Dale L. 1988. Articulation: A resolvable problem? In: *Shaping the future of foreign language education: FLES, articulation, and proficiency*. John F. Lalonde II, ed. Lincolnwood, Ill.: National Textbook Company. 11-31.

- Larson, Jerry W. 1989. S-CAPE: A Spanish computerized adaptive placement exam. In: Modern technology in foreign language education: Applications and projects. Wm. Flint Smith, ed. Lincolnwood, Ill.: National Textbook Company. 277-89.
- Larson, Jerry W. 1986. Testing in the language laboratory. In: Planning and using language learning centers. Jerry W. Larson, ed. Provo, Utah: CALICO, Brigham Young University.
- Medley, Frank W., Jr. 1985. Designing the proficiency-based curriculum. Proficiency, curriculum, articulation: The ties that bind. Alice Omaggio, ed. Middlebury, Vt.: Northeast Conference. 13-40.
- Rivera-La Scala, Gladys. 1989. The Annapolis interactive video project. In: Modern technology in foreign language education: Applications and projects. Wm. Flint Smith, ed. Lincolnwood, Ill.: National Textbook Company. 257-61.
- Rogers, Carmen Villegas, William H. Heflin, and John B. Romeiser. 1988. Conversion to a proficiency oriented curriculum at the university level. In: New challenges and opportunities. Ed. T. Bruce Fryer and Frank W. Medley, Jr.. Columbia, S.C.: Southern Conference on Language Teaching, University of South Carolina. 15-28.
- Rubin, Joan. 1989. LLD: The language learning disc. In: Modern technology in foreign language education: Applications and projects. Wm. Flint Smith, ed. Lincolnwood, Ill.: National Textbook Company. 269-75.
- Schulz, Renate A. 1988. Proficiency-based foreign language requirements: A plan for action. ADL Bulletin 19.24-28.
- Stansfield, Charles W., Dorry Kenyon, Fatima Doyle, Ricardo Paiva, and Ines Ulsh. 1988. Portuguese speaking test: Official test manual, 1988-89 edition. Washington, DC: Center for Applied Linguistics.
- Stansfield, Charles W. 1989. Personal communication.
- Stone, LeeAnn. 1988. Task-based activities: A communicative approach to language laboratory use. Iowa City, Iowa:IALL.
- Underwood, John H. 1987. Artificial intelligence and CALL. In: Modern media in foreign language education: Theory and implementation. Wm. Flint Smith, ed. Lincolnwood, Ill.: National Textbook Company. 197-225.
- Valdman, Albert. 1988. The symposium on the evaluation of foreign language proficiency: Challenges to the profession. New challenges and opportunities. Ed. T. Bruce Fryer and Frank W. Medley, Jr. Columbia, SC: Southern Conference on Language Teaching, University of South Carolina. 1-14.
- Verano, Miguel. 1989. USAFA interactive study in Spanish. In: Modern technology in foreign language education: Applications and projects. Wm. Flint Smith, ed. Lincolnwood, Ill.: National Textbook Company. 249-256.

The synergism of technology and theory in classroom second language acquisition research

Nina Garrett

University of Illinois at Urbana-Champaign

To postulate synergism between two phenomena or entities is to claim that together they can achieve more than is possible for either individually. My thesis in this paper is a dual one: first, that the use of technology to assist in classroom second language acquisition research can be enormously more significant if it is informed by a theoretical framework with which its capacities are inherently compatible, and second, that the most interesting and rewarding theoretical approaches to classroom second language acquisition research can only be adequately explored in a computer-based research instrument.

To support that thesis I must first spell out--in order to set them aside--the kinds of technology-based research on, or assessment of, classroom learning which I am not interested in exploring here.

First, I do not wish to explore here the use of the computer in methodological research evaluating the use of the computer in teaching language. Those of us who are convinced of its potential for enhancing or improving language learning are always being challenged to 'prove' it, to collect data in controlled experiments which will indicate whether or not using the computer is generally worthwhile. I do not believe that it is possible to carry out experiments which will give valid information on so broadly stated a question as 'Does technology use improve language learning?' though it is possible to design smaller-scale studies to explore one or another subpart of the question, as for example 'for what kind of learner, trying to learn what kind of language behavior, under what logistic circumstances, within what pedagogical framework, is what kind of computer-based material most likely to be helpful?'

Even more to the point is the fact that the use of computers does not constitute a method, approach, or philosophy, and therefore cannot of itself be evaluated as to its success as some kind of pedagogical principle independent of content, philosophy, or any number of situational and personal factors.

For present purposes, then, I am setting aside the use of the computer for collecting data with which to evaluate the effect of some pedagogical manipulation, and focusing instead on the computer's potential for supporting research on the acquisition process itself.

In principle any computer language lesson which elicits, collects, and 'analyzes' learner language data could be used in classroom second language acquisition research. However, just as many computer lessons are 'electronic workbook' exercises, where the pedagogical task is identical to that performed

in a workbook, research based on such computer programs would be 'electronic workbook' research, in which the computer substitutes exactly for paper and pencil in eliciting, collecting, and scoring learner data for us. We already have computer language materials which set the learner a task and 'analyze' the response--i.e. check for an exact match with the internally stored correct answer--at least for the purpose of giving 'right-wrong' feedback, but as yet few programs even collect those bare right-wrong scores for any use by teacher or learner after the lesson has been turned off.

It is not difficult to imagine language acquisition research projects which could make use of scores as data, *if* the task is such that a score (the record of a binary choice between a right and a wrong answer) confirms or disconfirms a relevant hypothesis about the acquisition process, or, to state it in slightly different terms, if the theoretical prediction being tested can be addressed by data quantifying the notion of mastery. That notion implies a view of language as a set of structures whose specification as discrete entities allows something theoretically meaningful to be said about them. A good example of such research would be the testing of parameter setting predictions generated by Government and Binding theory, where data on 'mastery' of certain structures are interpreted within a coherent, integrated, and systematic concept of language. A bad example of research using scores would be the Bilingual Syntax Measure, where the particular structures whose 'mastery' was assessed had no principled relationship to each other, and where the predictions being 'tested' did not derive from a coherent or integrated notion of language or language use. Both these research efforts can of course be undertaken without the computer, though it seems clear that much more data could be gathered and processed much more efficiently if the research instrument were 'computerized'. Nonetheless, in this kind of research theory and technology are still independent, forming no synergism.

When the computer is mentioned in connection with language research, it is often taken for granted that the connection is parsing, which is from one perspective the most sophisticated kind of linguistic analysis done by the computer. In parsing, algorithms based on artificial intelligence constructs are used to generate a model of the learner's utterance, or a model of the grammar underlying the utterance, and to compare that to a model of a corresponding but correct utterance in the target language. The output of this kind of computer analysis of interlanguage is not a score but a grammatically categorized description of the ways in which the learner utterance deviates from the norms of the TL. With all respect for the extremely interesting and significant work being done in this area, however, I suggest that in second language acquisition research our primary purpose is not simply to enumerate or to describe learners' errors but to understand why they make them.

Now, the distinction between scores, descriptions, and explanations is not clear-cut; in some cases a very fine articulation of a score will ipso facto serve as a kind of explanation of it. For example, if a learner gets a score of 70% correct on a quiz on a given grammar point, we know nothing about why she made the errors on the other 30%. If, however, she responded to all the items of one type incorrectly and all the others correctly, that explains her performance at least to some extent--she understands X but has trouble with

Y. But the crucial difference between a score and an explanation lies not in the degree of specification of the errors but in the explanatory power of the categorization of the language items--in the *theoretical* adequacy of the explanation. For example, suppose that in a language task dealing with the recipients of actions of giving and sending, a learner marks these indirect objects correctly in 85% of the obligatory cases and marks the rest as if they were direct objects. Without further information no explanation can be extrapolated. But if the correct forms are supplied wherever the indirect objects are persons and the errors are made where they are not ('He wrote the company an angry letter' or 'My aunt left the university all her money') then we can explain her performance as governed by an idiosyncratic rule that connects the formal marking for the indirect object incorrectly to the feature 'personhood' instead of the semantic role 'recipient'. And that explanation is theoretically interesting because personhood is a feature which governs a wide range of morphological markings in many languages; personhood is a psycholinguistic prime notion in the way human beings organize grammar. If a parsing algorithm--or a human researcher--is prepared to recognize semantic features such as personhood as features of a linguistic situation and categorize correct and incorrect usage on that basis among others, then scores or parses representing such patterns constitute explanations of language acquisition research data. It is the theoretical basis for interpreting the data, rather than the form of the data themselves, that establishes their significance.

But so far these lines of argument, all the examples I have cited, have related (1) the type of data--scores and descriptions--to (2) the theoretical framework within which they are interpreted, but have not related either to the way the data are collected. Scores can be collected, parsing analyses can be applied, in pencil-and-paper research. (In fact, computer parsers are still no match for humans.) Again, the computer can help collect, organize, and statistically analyze more data more efficiently, but that is synergism only in a trivial sense.

Furthermore, both scoring and parsing data, regardless of how they are generated, result from analysis of learners' language product, analysis of utterances after they have been made. But in SLA research we are interested not simply in a description of the product but in an explanation of it: we want to know why learner language or interlanguage is what it is. The why is important for two reasons, pedagogical and theoretical. *Pedagogically*, it may fairly be assumed that any activity which includes cognitive processing by adult learners can be improved by explicit assistance to the accuracy or appropriateness of that processing--that is, an explanation of how the processing is going wrong will help learners correct it. (This argument only works, of course, if we accept that language learning by the adult in the classroom has an intrinsic cognitive component which cannot be short-circuited--whether or not it is conscious or consciously addressed.) *Theoretically*, a substantive explanation of why a learner makes a certain kind of error can confirm or disconfirm hypotheses about classroom second language acquisition.

There are two major approaches to explaining second language acquisition. The approach developed within current Chomskyan theory explains SLA in terms of the abstract innate constraints which shape the

development of learners' grammar under various circumstances (relationship of L1 and L2, for example.) This approach tends to be rejected by language teachers, who see it as affirming the primacy of abstract structure over meaning in governing language learners' performance. In the approach to second language acquisition undertaken by language teachers the communication of meaning is seen as primary. (In fact, a high proportion of SLA work in ESL and applied linguistics explicitly rejects any focus on form, setting it in contrast to meaning.) If carefully constrained and related to each other in principled ways, both approaches can yield important insights into SLA and can actually complement each other.

If we pursue the latter approach to SLA, if language acquisition is regarded as the acquisition of the ability to convey meaning, research on CSLA must explore the ways learners develop that ability--and that necessitates a research focus not on the *product*, the learner's utterance, but on the ongoing construction of meaning--on the processing of language. (Attempts to derive explanations of how learners construct utterances from the analysis of the final product is rather like trying to reconstruct the directions for concocting a complicated casserole from a chemical analysis of the food as it is set on the table.)

Let me digress a moment to make explicit a distinction which is sometimes blurred in discussions of SLA, that between the process of acquisition and the processing of language. Acquisition is a process which takes place over time. Research on acquisition must therefore either consist of a longitudinal study, i.e. several related studies over time, or some valid substitute for it. But an exploration of learner language at any given point in time, any one of the several studies that make up a longitudinal set, is a study either of the language produced at that time--the product--or a study of the mental processing, the encoding of meaning into language or the decoding of language into meaning. The 'process' (i.e. of language acquisition) can only be studied indirectly, through successive analyses of either product or processing. The product can be analyzed in terms of the rules of the system (linguistic, sociolinguistic, whatever aspect of the system is applicable) which describes the target language, and the result of the analysis is then an account of the discrepancies between the descriptive rules of the TL and the related descriptive rules of the interlanguage.

But how do we investigate the ongoing processing of meaning? What kind of data, and what kind of data analysis, will give us insight into learners' idiosyncratic processing? We could in principle ask learners what they were thinking, why they chose certain words or structures to express themselves, but it is obvious that doing so would be so disruptive of any communicative attempt as to vitiate the results entirely. The computer offers an unprecedented research methodology.

Implementing this methodology requires a major rethinking of the premises on which most computer-based language materials have been developed until recently. Most lessons have presented and/or drilled facts about language--vocabulary or grammar--so that learners could master these outside of class, freeing the teacher and the class hour for communicative activities. (Some 'communicative' lessons have offered games, simulations, or problem-solving activities which depend on comprehension or require only

minimal language production not evaluated for accuracy. In principle such programs could be designed so that data on learners' engagement with them provided evidence of language processing, but I will limit myself here to discussing more text-oriented materials.) If the technology, instead of serving as a utility for drilling linguistic facts, provides an environment which supports and assists learners in their attempts to express meaning in the target language and also includes a facility for keeping track of how learners do so, it can elicit, collect, and analyze hitherto unattainable data for classroom second language acquisition research.

For example, to study how learners actually set about reading in a second language, a computer-based reading comprehension task could be created out of any text with a wide range of help material accessible through keypresses --glosses at several levels of difficulty and for various purposes, such as straight lexical meaning, citation forms, analysis of grammatical form, variations in register or usage, contrasts between literal and idiomatic meaning, any of these in the native or target language. A variety of strategies could be suggested to assist in top-down or bottom-up processing--highlighting techniques, markup with symbols to cue certain kinds of reading, timed presentation--any or all of these, as controlled by theoretically motivated hypotheses. And a computer program running invisibly behind the task could keep track of precisely what each learner calls for at precisely which point in working though the passage and answering questions on it. These data could then be examined for correlation with a variety of independently assessed theoretically motivated characteristics of the learner or the nature of the task. For example, do learners of a certain age, or at a certain level of learning, or with a certain cognitive style, or in reading certain kinds of material, tend to use particular strategies, or interpret or misinterpret the material in predictable ways? How consistently? On what basis? Are they aware of syntactic clues, even when these have not been explicitly discussed? Do they use L1 strategies for inferring topic?

Analogous research can easily be imagined for listening comprehension, when the learner is presented with audio or video material under the computer's control. Given a controlled variety of helps--audio replays of large or small segments, appearance on screen of single words, whole phrases, glosses, paraphrases, grammatical analyses, graphics, still pictures, video to provide the context of facial clues--what path do learners of particular characteristics follow, what strategies do they use, and with what success in comprehension?

Research on vocabulary use and acquisition suggests the structuring of a multimedia (especially a hypermedia) environment allowing audio, video, or graphic representation, copying, repetition, examples, choice of native or target language definitions or paraphrases, to help learners construct semantic networks. What kind of networks do they construct?

In processing vocabulary or reading or listening comprehension, learners are dealing with preestablished meaning and are learning to structure their own form-meaning connections to match those of TL native speakers. Since the meaning is controlled by the task, the researcher can not only track the learners' moves but can also analyze the evidence of comprehension, give feedback on the basis of the analysis, and keep track of what they do on the

basis of the feedback. If errors (in comprehension, in grammatical judgments, in syntax, in discourse organization, etc.) are indicated without explanation, can learners correct them? Can they correct without understanding, and if so, what does that indicate about the processing? What assumptions do learners of different characteristics make about their own errors? What are their short-cuts, their working hypotheses about how language is learned? Is it possible to tell the difference between careless errors and those representing systematic problems with processing?

Research on learners' ability to express themselves in writing becomes possible with the introduction of a class of programs like Système-D, where learners are assigned to write a French composition and have access to a lexicon with usage examples and notes, a reference grammar including all verb forms, idiomatic vocabulary and phrases covering a wide variety of topics and conversational functions. With an attached utility recording the learner's every keypress, a researcher can track what helps are called for at what point in the writing, and what the learner does with the information looked up. For example, do independent measures of impulsivity/reflectiveness predict differences in the way learners edit their writing? Do the strategies change at advanced levels of learning?

In setting spontaneous writing tasks we cannot know *a priori* what meaning learners are trying to express, and current parsers are not able to analyze learner input for anything more complex than simple morphological and syntactic patterns. But it is easy to constrain the meaning learners are asked to convey by asking for descriptions of pictures or video, paraphrase, translations, etc., so that the computer 'knows' the meaning to be expressed and can be programmed to analyze learners' idiosyncratic form-meaning connections without making learners aware that any particular structure is under scrutiny. To pick up on the example used earlier, if the hypothesis is that learners tend to use indirect object markers only on person-objects rather than to mark recipienthood, the task can be structured to include both personal and impersonal indirect objects and data can be collected not only on the percentage of error in each category but also on what interpretation learners place on error markup--do they know what is wrong and how to correct it? Do they try first to change roots, or endings, if they don't know why it's wrong?

These suggestions may meet with arguments that learners' interaction with computers is not truly communicative, or authentic, because computers do not communicate like human beings. Obviously, the technology cannot assess learners' ability to handle *sociolinguistic* variation. Nonetheless, in the notion of communicative competence that inheres in applied linguistics approaches to SLA, the pendulum has swung so far towards the domination by *sociolinguistic* factors that we are in danger of overlooking the fact that there are a lot of other kinds of meaning which need to be conveyed in any utterance. Many of the semantic concepts that were organized by David Wilkins in his discussion of notions (the disastrously overlooked part of the notional-functional syllabus) are basic to the communication of meaning and underlie all the *sociolinguistic*, pragmatic, or discourse features which figure so prominently in communicative competence. These features of inter-

language can be explored authentically in written language on the computer, as well as in spontaneous oral discourse.

The significance of research on learner language processing supported by this kind of computer use is limited only by the theoretical rigor of the hypotheses we develop to explore with it, and is in fact far greater than is commonly recognized. Insofar as they think of using computers at all, most language teachers think in terms of drills and linguists think in terms of parsing and artificial intelligence techniques. Continued work in these areas will certainly provide us with undreamed-of power, but substantive work in SLA need not wait for future developments; the technology is already far more sophisticated than our ability to make theoretically motivated use of it. My point here is that without a theory that postulates an active role for language learners in organizing, processing, and generating linguistic, sociolinguistic, etc. knowledge, we are likely to continue to use the computer only in relatively superficial ways to elicit, collect, and analyze data on the product of learners' language behavior; we cannot undertake research on processing without a theoretical concept that makes testable predictions about processing. (We can, of course, undertake open-ended pilot studies, to help us generate testable hypotheses.) Conversely, theories about second language learners' processing are not testable without a research methodology which allows immediate ongoing objective collection and analysis of learners' processing, their attempts to come to terms directly with meaning in another language than their own.

Charles Ferguson's presentation at this Round Table persuasively urged the use of second language learners' data to test language acquisition hypotheses generated by linguistic theory, citing the need to account for language-specific phenomena as well as acquisitional universals. The use of technology to elicit, collect, and analyze interlanguage data would allow the often overloaded language teacher to undertake far more ambitious second language acquisition research than has ever before been possible, in the course of delivering innovative and pedagogically sound language learning assistance. That is to say, CALL has a natural mate in CARLA (computer-assisted language acquisition research), and the synergism between them can have unprecedented benefits for foreign language pedagogy, for second language acquisition theory, and for the careers of language teacher/researchers.

References

- Ferguson, Charles. 1989. Language teaching and theories of language. In: Georgetown University Round Table on Languages and Linguistics 1989. Washington, D.C.: Georgetown University Press.
Wilkins, David. 1976. Notional Syllabuses. Oxford: Oxford University Press.

Semantic subclasses of temporal nouns

Michael Zarechnak
Georgetown University

Abstract. The purpose of this paper is to illustrate and justify semantic coding of temporal nouns in Russian for Russian-to-English Machine Translation to facilitate syntactic parsing of Russian sentences and their subsequent translation into English. The paper deals only with the semantic classification, based on A.A. Kholodovich's theory of subclasses. The temporal nouns are selected for justification of four subclasses: cyclic, supercyclic, subcyclic, and noncyclic. The codes are to be stored in the dictionary. The parsing rules based on the cooccurrence restrictions between these four subgroups facilitate syntagmatic bracketing of noun groups as an input for the transfer and synthesis into English.

The purpose of this paper is to group temporal nouns into subgroups and analyze the cooccurrence of privileges and restrictions among the temporal subgroups. These restrictions are given within the Kholodovich concept of the nucleus and its optimal environment. It is assumed that the reader is acquainted (1) with the various ways a temporal noun can be identified, and (2) the main temporal figurae assigned to temporal adverbs (see Appendix). We shall use morphological, syntactic, transformational, and pure semantic criteria for subclassification of temporal nouns. In the subclassification proposed here we shall use three main criteria: (1) the concept of primary vs. secondary, (2) the concept of a cycle vs. noncycle, and (3) the concept of subcycle and supercycle. When we take such words as *god* 'year' and *noch* 'night', we can contrastively observe the following features. The Russian word *god* 'year' indicates a one-cycle duration, during which the earth revolves once around the sun. In this word there is no additional concomitant semantic feature. We shall refer to such a noun as a 'cyclic temporal noun'. In *noch* 'night', on the other hand, we have less than a cycle, and besides we have an additional semantic feature, namely, that the sunlight is not seen. We shall refer to such a noun as a 'subcyclic secondary temporal noun'. The opposite of the cyclic and subcyclic nouns will produce what we call 'noncyclic' and 'supercyclic nouns', for example, two words such as *vechnost'* 'eternity' and *molodost'* 'youth'. In the first we have no cycles to talk about. Such a noun has no boundaries. The word *molodost'* involves more than one cycle. It refers to a particular stage in the development of a living creature, especially a human being. The type *vechnost'* is a 'noncyclic primary temporal noun'. The type *molodost'* 'youth' is a 'supercyclic secondary temporal noun'. If we put the above-mentioned criteria (cycle and subcycle) on a matrix, we shall see the distribution displayed in Figure 1.

We may now try to analyze them within the framework of the N1 of N2 noun government structure. We will examine the relations between the

temporal nouns themselves and the relations between the temporal and nontemporal nouns. It may happen that we shall see an utterance like *utro* (N1) *voskresnogo dnja* (N2) 'the morning (N1) of the Sunday (N2)' where both the N1 and N2 positions are occupied by temporal nouns. In such a case we should examine the cooccurrence privileges between any pair of the these subclasses. On the other hand, if N1 or N2 is occupied by a nontemporal noun, we would like to know what nouns occur in the vicinity of the temporal noun.

Figure 1. Temporal nouns.

Criteria:	Primary		Secondary	
Types of nouns:	cyclic	noncyclic	subcyclic	supercyclic
(1) <i>god</i> 'year'	#			
(2) <i>vechnost'</i> 'eternity'		#		
(3) <i>noch'</i> 'night'			#	
(4) <i>molodost'</i> 'youth'				#

Nontemporal nouns cooccurring with temporal nouns. To increase control over our operations we divide the cooccurring and the temporal nouns into the following further subgroups (on the morphological level): (1) nouns occurring only in the plural, (2) animate and inanimate, (3) deverbal abstracts (such as *pisanie* 'writing') vs. deverbal concrete (such as *pis'mo* 'letter'), and (4) quantifiers and qualifiers.

Nontemporal nouns occurring in position of N1, when the N2 is a temporal noun which occurs only in the singular or only in the plural. We change singular to plural, and vice versa, in order to see whether the time nouns change their meaning. These nouns were found in *Russkij Les* 'Russian Forest', a novel by Leonov. Our observations are checked in *Ushakov's Dictionary* (a sort of Russian Webster's).

Singular only: *ujma* 'a lot'

Plural only: *sumerki* 'twilight', *sutki* '24 hours (day and night), *budni* 'weekdays' (Leonov 1954), ...*u russikh eshche est' ujma* (N1 of N1 sg.) *bespechnogo vremeni* (N2 time-noun, sg.) 'Russians still have a lot of carefree time'. Here it is impossible to change the singular N1 into plural. (See also *Ushakov III*, 912).

The form *ujma* has only one meaning: *ochen' bol'shoe kolichestvo* 'a very large quantity'. We classify *umja* as a singular only on the morphological level, and as a 'quantifier' on the semantic level. (1) The syntactic property of a quantifying noun is that it cannot stand alone. (2) A substitution test using *mnogo* 'much', *malo* 'many', *kakoe-to* 'some', *kolichestvo* 'certain

quantity' corroborates lexically the fact that *ujma* belongs in the quantifying subclass. If a time-noun is singular, it can express any arbitrary length of objective or subjective time; but once the singular form is defined as a certain arbitrary length of time, then, if the noun is used in the plural, each segment must be the same: e.g. *vremena goda* 'seasons of the year', *zima* 'winter', *vesna* 'spring', *leto* 'summer', *osen'* 'autumn'. Thus, we can compile a list of singular nouns which cooccur with time-nouns.

(1) Quantifiers:

- (1.1) *poltora*, *poltory* 'one and a half'
- (1.2) *dva* (m.), *dve* (f.) 'two', *tri* 'three', *chetyre* 'four'
- (1.3) *chetvert'* 'a quarter' (*dobriju* 'good', *upushchennuju* 'lost', *kazhduju* 'each'), *dol'ka* 'part', *chast'* 'segment', *ostatok* 'remainder', *ujma* 'a lot'

(2) Qualifiers: *blizost'* 'proximity', *gorech'* 'sadness', *blagost'* 'grace'

- (3) Compound prepositions: *za davniost'ju* 'on the grounds of past events', *po proshestvii* 'after', *na iskhode* 'towards the end of', *ot nachala* 'from the beginning', *na protjazhenii* 'during', *v preddverii* 'in advance of', *v svete* 'in light of', *v nachale* 'in the beginning', *v konce* 'at the end'

(4) Unitary nouns: *solnyshko* 'little sun', *solnce* 'sun'

- (5) Inanimate nouns: *svet* 'light', *nebo* 'sky', *vozdukh* 'air', *tishina* 'silence', *veter* 'wind'

- (6) Deverbal nouns: *likovanie* 'joy', *razvitie* 'development', *khod* 'move', *razmaki* 'scope', *ozhidanie* 'expectation', *obstanovka* 'circumstance', *merka* 'measure', *blagovest'* 'good news', *spasenie* 'saving', *uklad* 'structure', *nepogoda* 'weather'

- (7) Temporal nouns cooccurring with temporal nouns: *rassvet* 'dawn', *starina* 'antiquity', *nachalo* 'beginning', *budushchee* 'future', *proshloe* 'past'

These nouns deserve additional comments. The conclusions we shall make are preliminary. Further statistical studies will be necessary to confirm or contradict our findings.

We can say that despite various morphological and syntactic features characteristic of the nouns in group (1), semantically there is no difference among the quantifying words in the sense that any quantified word is related to its quantifier through a 'suppressed predicate', i.e. when we say *pjat' dnej*, we mean really 'there are five days'--yet this predicate is easily predictable.

Nouns from group (2), being deadjectival, are capable of playing a role in the compound predicate within the kernel type 'N is A', where A (adjective) is derived from such a noun. Compare *blizost' oseni* 'nearness of autumn' ==> *osen'* *blizka* 'autumn is near'-- with the zero auxiliary copula.

Nouns from group (3), nouns governed by prepositions, are used contextually only as 'temporal compound prepositions' (TCPs). Here it suffices to notice that temporal nouns participating in TCPs must be in the singular, and cannot be used in that function in the plural.

In group (4), unitary nouns, the word *solnyshko* can also have a secondary temporal meaning. Logically, it belongs to a unit subclass by itself. It cannot

be used in the plural. Compare: *do rassveta* vs. *do solnyshka*, where in both utterances the message is the same, i.e. 'before the sun rises'.

Nouns from group (5) are not used in the plural with temporal nouns.

Group (6), deverbal nouns, are consistently used only in the singular in the text. The deverbal nouns are used for transformational tests of the type N1 of N2 ==> NV, where V comes from N1, while N2 is occupied by a temporal noun.

Comparing a group of nouns from (1), (2), and (6), we may observe that in (1) the predicate is suppressed, in (2) it is semisuppressed, and in (6) it is explicitly present.

Group (7) contains time-nouns with deictic functions (1), i.e. pronoun-like nouns referring to other nouns.

In conclusion, we may say that when a word is used in its peripheral meaning (*veter* meaning not 'wind', but 'tendency') then, obviously, the environment is responsible for the singular usage since *epokha* 'epoch', *vremja* 'time', normally do not allow for the cooccurrence of *veter* with its primary meaning, i.e., 'wind'.

Animate nouns cooccurring with temporal nouns. We now turn our attention to the problem of 'animate' nouns cooccurring with primary temporal nouns. Consider the following:

- (1) *aviator* (N1) *svoego vremeni* (N2) 'aviator (pilot) of his time'
- (2) *starozhily* (N1) *tekh vremen* (N2) 'old-timers of those years (times)'
- (3) *zhitel'* (N1) *toj pory* (N2) 'resident of that season (time)'
- (4) *ljudi* (N1) *togo vremeni* (N2) 'people of that time'

The transformations for these 'N1 of N2' structures show that some of them are easier to change into subject-predicate than others. Thus, (1), *aviator, zhivshij, letavshij v svoe vremya* 'an aviator who lived, flew in his time' shows that N2 is an adverbial (D) modifier and accordingly a predicate has to be added (the underscored words).

In example (2), *starozhily tekhn vremen* ==> *ljudi zhivshie v te vremena i zhivushchie tam i teper'* 'who lived at that time and are still living there now', note that the predicate is added from the part of the *starozhily* which is composed of *staro* + *zhil-y*. The form *zhily*, by itself, does not occur.

Example (3) shows the following transformation: *zhitel' toj pory* 'resident of that time' ==> *ljudi, kotorye zhili v tu poru* 'people who lived at that time'.

Example (4) yields the following: *ljudi togo vremeni* 'people of that time' ==> *ljudi, zhivshie v to vremya* 'people who lived at that time'.

Thus, a feature common to any 'N1 of N2', where N2 is a temporal noun, is the fact that in transformations N2 cannot serve as a subject for N1.

On the other hand, when N1 is an animate and/or a deverbal concrete noun (examples 1, 2, and 3), it functions as a subject, and as defined in this section, we expect to get an interpretation for a predicate. We produce this by inserting a plausible predicate. This fact automatically changes the function of N2 into an adverb (D). The kernelization (making a simple sentence: 'subject-predicate + adverb' from 'N1 of N2') takes the shape of the following transformation: N1 (anim.) of N2 (temp.) ==> N1-subject + RP (Relational

Predicate) + D <= N2. In short, a minimal group N1 of N2 ==> NVD (<= N2).

Inanimate nouns cooccurring with temporal nouns. These nouns (be they morphologically nonverbal, concrete, or deverbal concrete, such as *zamok* ‘castle’ vs. *zapiska* ‘note’, respectively), if governing N2, function as subjects with a suppressed predicate, and N2 functions as an adverbial modifier of time. Consider the following:

- (1) *zamok 17-go stoletija*, i.e. ‘built in the 17th century’, *postroen* ‘built’
- (2) *gazety togo vremeni* ‘newspapers’, i.e. ‘retained from that time’, *sokhranivshiesja* ‘preserved’
- (3) *zapiska 17-go goda* ‘note’, i.e. ‘pertaining to the 17th year’, *otnosjashchajasja* ‘relating’
- (4) *leto 17-go goda* ‘summer’ i.e. ‘which was in the 17th year’, *byvshee*

The words on the right suggest a plausible (but not unique) predicate by which a sentence can be formed from these ‘N1 of N2’ structures.

Deverbal nouns cooccurring with the temporal nouns. If N1 is a deverbal process noun, then the relation of N1 to N2 is either:

- (1) a ‘predicate-subject’ relation as seen in the transform where: N1 ==> ‘predicate’ e.g. *nastuplenie uchebnogo goda* ‘arrival of the school year’ ==> *uchebnyj god nastupaet* ‘the school year arrived’ and N2 ==> ‘subject’, e.g. *konec sentjabrja* ‘end of September’ ==> *sentjabr’ konchaetsja* ‘September draws to an end’, *Nachalo goda* ‘the beginning of the year’ ==> *god nachinaetsja* ‘the year begins’

or

- (2) a ‘predicate-object’ relation as seen in the transform where N1 ==> ‘predicate’ and N2 ==> ‘object’ as in: *chtenie knigi* ‘reading of a book’ ==> *knigu chitatut* ‘the book is read’ and *ozhidaniie zimy* ‘the expectation of winter’ ==> *zimu zhdut* ‘the winter is expected’ or *ozhidanie zimy* ==> *ozhidajut zimy*.

Thus the word *kniga* and the two-place predicate (*Kto?* ‘Who?’ *Chto?* ‘What?’) *chitat’* ‘to read’ are in an antisymmetric relation to each other.

N1 is a temporal noun and N2 is a deverbal noun. N2 may be a deverbal noun. Consider: *gody lishenij* ‘years of deprivation’. The relation is:

N1 of N2 ==> PN = D (<= N1) RP (Relational Predicate) plus N2.

Thus, *lishenija imeli mesto v... godakh* (PN = D) ‘deprivation took place in ... years’ or *v...godakh* (PN = D + C), *kogda lishenija* (N2) *imeli mesto* (RP) ‘in...years, when deprivation took place’. N2 could also be a deadjectival noun, such as *vremja* (N1) *zhestokosti* (N2) ‘time of cruelty’. Here again a

restored predicate is necessary. The transformation is: N1 of N2 ==> D (< == N1) RP + N2. *Vremja* (N1), *v techenie kotorogo kto-to byl zhestok (konchilos)* 'the time during which someone was cruel (ended)'.

It is of interest to note that if N1 is occupied by a temporal noun and N2 by a nontemporal noun, N2 and N1 both serve as 'subjects', and the structure must be so modified as to show it in transformation. This requirement involves, of course, a heterogenous level in the kernelization of this 'N1 of N2', i.e. such conjunctions as *kogda* 'when' have to be used. To put it differently, when N1 is a temporal noun, as here, its kernelization produces a complex sentence, since 'N1 and N2' serve as the 'subject', and the 'predicate' must still be added from the wider context. Such a predicate has been indicated by parentheses (*konchilos*' 'ended').

N1 is a temporal noun and N2 is a quantifying or qualifying noun. On the basis of the observation that N1 can be an adverb, a predicate, and a subject when N2 is a temporal noun, we can on the 'content' level assume that when N1 is an adverb, it is a modifier of the restored predicate. The particular nature of this modifier depends on the semantic structure of the stem of N1. Then we seek the various possible semantic structures of temporal nouns in the N1 position so that we may accordingly classify them into subgroups.

Relations between temporal nouns cooccurring in both N1 and N2 positions. The four subgroups of temporal nouns established at the beginning of this paper (cf. Figure 1) can cooccur with one another, with certain restrictions. I shall label each of these subgroups of temporal nouns by the abbreviation listed after the name of each subgroup: cyclic (cyc.), subcyclic (sbc.), supercyclic (spc.), and noncyclic (ncy.). As previously stated, temporal nouns can be objectively and conveniently classified into four subgroups:

cyc. cyclic: e.g. *god* 'year', *mesjac* 'month', *sutki* '24 hours'

sbc. subcyclic: e.g. *den'* 'day', *noch'* 'night', *semernki* 'twilight'

spc. supercyclic: e.g. *molodost'* 'youth', *starost'* 'old age', *molodezh'* 'young people', *vozrast'* 'age'

ncy. noncyclic: e.g. *vremja* 'time', *sovremennost'* 'contemporary', *pora* 'time', *srok* 'time, period, deadline', *zhizn'* 'life'

The table displayed in Figure 2 is very close to one's own intuitive notions. Indeed, the minus in cyc./sbc. is given on the basis of 'part vs. whole' within cyclic time units, such as *den'* vs. *sutki* vs. *mesjac*. The minus in cyc./ncy. is given on the basis of a polar contrast between the 'segmental' vs. 'nonsegmental' temporal nouns such as *god* vs. *vremja*. The minus in spc./sbc. is given on the basis of a secondary temporal noun vs. a part of a cyclic temporal noun which is a non sequitur, such as **molodezh'* *sroka* 'youth of a deadline'.

The minuses in ncy./sbc. and ncy./spc. are caused by logical contradictions between the noncyclic (ncy.) vs. the cyclic (cyc., sbc.), and primary temporal (ncy.) vs. secondary temporal (spc.).

Figure 2. Table of relations for N1 and N2.

N1	N2: cyc.	sbc.	spc.	ncy.
cyc.	N1 = D N1 N2 +	nest	N2 = subj. N2 - subj.	nest
sbc.	N1 = D N1 N2 +	N1 N2 +	N1 = D N2 - subj.	N2 figur.
spc.	N1 = subj. N2 = D +	nest	N1 = subj. N2 = obj.	N2 = subj. N1 = pred. +
ncy.	- +	- +	-	N2 figur. +

If we reverse the order of the above expressions, there is no correspondence in terms of the presence of a minus except with the pair ncy./cyc. and cyc./ncy.:

- (1) a minus for cyc./sbc., but no minus for sbc./cyc.
- (2) a minus for spc./sbc., but no minus for sbc./spc.
- (3) a minus for ncy./sbc., ncy./spc., but no minus for sbc./ncy., spc./ncy.
- (4) a minus for cyc./ncy., but no minus for ncy./cyc.

Now we can see that the noncyclic are opposed only to the secondary temporal noun, i.e. we can use spc./ncy.: *molodezh' sovremennosti* 'youth of contemporary time', but we cannot say ncy./spc.: **sovremennost'molodezhi* 'contemporariness of youth'.

It is obvious that larger statistical studies are needed. When an N1 is governed by a preposition, it has to be checked for participating in a CTP (Compound Temporary Preposition), as has been indicated.

Summary of the functional roles of 'N1 and N2'. 'N1 of N2' may constitute part of a 'PN1 of N2' structure. In this case we have several possible varieties.

Within 'PN1 of N2', PN1 may function as P, O, A, or the nominal part of a compound predicate.

We may also have PN1 without N2 where N1 is a time-noun, e.g. P-PN: *i sto semnadcat' tysjach v god* (Ushakov 266), where PN is a plain PN without the N2. It functions as a preposition with the meaning *vo vremja, za* 'during, in the course of, in'.

The plain PN can also be used in reversed order and has the meaning of the approximate time used for some action. Consider: Leonov 1954:12 *nedeli cherez tri* (Leonov 1954:12) 'in approximately three years', *goda na dva* (Leonov 1954:59) 'for approximately two years'; PN = CTP: v prodolzhenie

posledujushchej chetverti veka (Leonov 1954:304) 'during the subsequent quarter of a century', *po davnosti (istekshikh) let* (Leonov 1954:35) 'on grounds of past years', *v rascvete (tvorcheskikh) sil* (Leonov 1954:651) 'in the prime of one's life, in blossom of creative years'.

PN = D. If PN is spelled together, it functions as an adverb (D). Consider: *let vosem' podjad* (= *Pod* + *rjad*) 'approximately eight years', *let vosem' sriadu* (= *S* + *rjad* + *u*) (Leonov 1954:11) 'straight running in succession', *18 let nazad* (= *na* + *zad*) (Leonov 1954:30) '18 years ago'.

Let us recall that in the section devoted to the classification of adverbs, we established among other subgroups, the subgroups of 'continuity' and 'before'. The examples containing *podjad* 'straight', and *nazad* 'back', cooccurring with temporal nouns, represent these two subgroups, respectively.

If these two functional roles are not played by the N1, then we can add another factor to identify its role: the presence or absence of an 'adjectival modifier' in front of N1, or N2, or both. Here I shall mention only the main factors involved in an 'AN' structure.

The list of adjectival forms (adjective, pronominal, participle) which can precede either of the nouns follows: *poslednij* 'last', *blizhajshij* 'next', *neminuemij* 'unavoidable', *vtoroj* 'second', *tot* 'that', *zhestochajshij* 'cruellest', *cel'nyj* 'whole'.

Some of these adjectival modifiers are optional while others are mandatory. Here are a few examples of A which cannot be dropped: *skuku (voennykh) budnej* (Leonov 1954:405), vs. an ommissible A in *v kurse (vsekh) sobytij* (Leonov 1954:298). The reasons for the omission of adjectives are not given here.

The adjective with N1: (*celaja*) *noch'* (*skazochnykh*) *prikluchenij* 'the whole night of fantastic adventures' (Leonov 1954: 608); *cel'nykh vosemnadcat'* *let* 'straight 18 years' (Leonov 1954:99), *poslednie desyat'* *let* 'last 10 years' (Leonov 1954:406).

The adjectives with N1 and N2: *v (samom) kotle (zhestochajshikh) sobytij* 'in the very boiler of the most cruel events' (Leonov 1954:352); (*bylovoj*) *uklad (nashikh) predkov* '(life) style (of our) forefathers' (Leonov 1954:251). An adjectival modifier may be semantically as important as the noun it modifies. Syntactically, an adjective may occur as the subject, object, or an adverb for the noun it modifies. Compare: *ego delo* 'his affair', *pit'evaja voda* 'drinking water', *vcherashnij den'* 'the yesterday's day', respectively.

Appendix: Semantic temporal figurae.

Words Figurae:	1	2	3	4	5	6	7
1 <i>vesnoj</i> 'in springtime'	+	-	-	-	-	-	-
2 <i>zavtra</i> 'tomorrow'	-	+	-	-	-	-	-
3 <i>davno</i> 'long ago'	-	-	+	-	-	-	-
4 <i>do groba</i> 'until deal'	-	-	-	+	-	-	-
5 <i>bystro</i> 'quickly'	-	-	-	-	+	-	-
6 <i>chasto</i> 'frequently'	-	-	-	-	-	+	-
7 <i>bezotryno</i> 'always'	-	-	-	-	-	-	+

1. A figure of duration which is a single, natural unit.
2. A figure of duration which consists of a set of objectively provided temporal units.
3. A figure of duration where both length and assignment to the specific temporal segment is done by man, i.e. the speaker.
4. A figure of boundaries of time.
5. A figure of speed.
6. A figure of frequency.
7. A figure of eternity.

References

- Hutchins, W.J. 1986. Machine translation: Past, present, future. New York: John Wiley. 37, 71 ff., 133-36, 161, 220.
- Leonov, Leonid. 1954. Russkij les'. Moskva: Molodaja Guardija.
- Weinrich, Uriel. 1963. On the semantic structure of language. In: Universals of language, ed. J.H. Greenberg. Cambridge, Mass.: MIT Press. 124 and 2.2.2.
- Zarechnak, M. 1967. Application of A.A. Kholodovich's theory of subclasses to Russian temporal nouns. Harvard University doctoral dissertation. 145-68, 78-81.
- Zarechnak, M. 1979. The history of machine translation. In: Machine translation (Trends in linguistics: Studies and Monographs no. 11), by Bozena Henisz-Dostert, R. Ross Macdonald, and Michael Zarechnak. New York/The Hague: Mouton. 1-87.

From wire recorder to satellite dish: The impact of technology on language teaching

Protase E. Woodford
Educational Testing Service

When I set down the title for my remarks today, I must have been taking the short view. Long before the wire recorder language teachers and learners must have used mechanical and technical aids. Champollion, in 1799, used the Rosetta stone to decipher the hieroglyphs. The stone itself needs be credited to the priests of Ptolemy V in the second century B.C. It is natural to believe that we are innovative and creative, that our particular skills and abilities result in great breakthroughs. For the last decade we have been focusing on 'proficiency'. The Foreign Service Institute language proficiency interview that served as a catalyst for subsequent work on 'language proficiency' was, indeed, a breakthrough, a performance test for foreign language. But, as you know, the ancient Israelites used a one-item language proficiency test when they required that any suspected Gileadite pronounce *Shibboleth*...

And he said *Sibboleth* for he could not frame to pronounce it right. Then they took him and slew him at the fords of the Jordan, and there fell at that time forty and two thousand.

The ancient Israeliite certified examiners also knew and applied the concept of the cut-score.

We live in different worlds. Some of our worlds overlap. Most of us here are language teachers, potential, practicing or past. Although we belong to that world we also inhabit smaller, more tightly delineated worlds. We are teachers of Chinese, or French. We are teachers of small children, adolescents, young adults, old folks. We are applied linguists and sociolinguists, methodologists and medievalists. Many of us do not teach anyone, directly, but rather through what we write or film or record. We are conditioned by the cultures of our worlds and, therefore, share their visions and hopes, their fears and prejudices. We also belong to even larger worlds; we are Asians or Africans, Europeans or Americans, Anglophones or Francophones.

I am here today because in one of the worlds I inhabit there is James Alatis. Jim is a friend. He asked me to speak. I do not like to say no to friends. But the world of the *Georgetown University Round Table on Languages and Linguistics* is not a world in which I feel part of the native fauna. It has seemed to me a place where great intellects have gathered to discuss esoterica, issues understood only by the initiated and not by the general body of disciples. I am, by training and experience, essentially a school teacher, a teacher of Spanish, although other activities have taken up

much of my time. The theme of this year's Round Table is one that should be of great interest to a Spanish teacher: Teaching, testing, and technology. And it is.

Last year I had the privilege of being at the United States Naval Academy in Annapolis to collaborate on an interactive video program designed to develop the listening skills of midshipmen in Spanish. (By the way, midshipmen are midshipmen whether male or female.) My specific responsibility was to contribute to the evaluation of the interactive video program. It was easy to get me to go there. I am an evaluator, the use of interactive video in language teaching is truly exciting; and, the hook, I was able to get back into the classroom and teach. It also became very clear to me that teaching a dozen midshipmen, highly motivated, very bright, was not nearly the challenge that teaching 30 or so seventh graders at State Street Junior High School was. The major difference was that in 50 minutes with my mids, I actually taught for 48 or 49 minutes--'Lessons from the past with a view toward the future.' After more than 30 years in this profession I feel that major concerns and issues simply recycle themselves every decade or so--Teaching, testing technology. Let's start in the middle and work to the ends.

The way we test has reflected how we teach. Look at the college entrance examination board foreign language tests of the 1930s, 1940s, and 1950s, and the cooperative classroom tests of that era, then the only nationally available foreign language tests. The college board French test of 1931 consisted of five parts:

- (A1) Passage to be read for understanding.
- (A2) Passage to be translated into English.
- (B) Prose composition English-French translation.
- (C) Essay: Free essay in French of 150 words.
- (D) 15 short questions.

The test books were scored by hand on a scale of 0-100. The test took one hour. There were different tests for different levels--French II, III, IV.

It is interesting to compare the volumes for the different language tests across time. In 1938, 5,952 examiners took the French examination, 4,113 sat for the Latin examination; 1,093 for German, 153 for classical Greek, and only 351 for Spanish.

In 1966, the year I joined the Educational Testing Service, the volumes were: French, 86,581; Latin, 20,670; German, 14,804; Greek, 152; and Spanish, 39,502. In 1966, tests were also offered in Hebrew and in Russian, with volumes of 837 and 1,273, respectively.

In 1988, the volumes for those language were: French, 24,848; Latin, 3,600; German, 3,512; Hebrew, 456; and Spanish, 30,637.

If we accept that testing tends to reflect how and what we teach, then the volume trends in the advanced placement (AP) program should be informative.

The modern language tests in the AP program include measures of listening, writing, and speaking, as well as reading. The original AP French and Spanish examinations were literature examinations. In the early seventies an alternative AP 'language' examination was introduced for Spanish and, later,

for French. In 1978, 2,958 students took the French literature test, and 1,396 the language test. In 1988, 9,277 students sat for 'language' and only 1,996 for 'literature'. In Spanish, in 1978, 2,025 students sat for the 'language' test and 1,534 for 'literature'; in 1988, 16,227 for 'language' and only 2,614 for 'literature'. Note, too, that the total number of French candidates in 1978 was 3,994, total Spanish, 3,553; French enjoying about 12% greater volume. By 1988, the total French volume had risen to 11,243, but Spanish had reached 18,841; whereas the numbers of test takers in the regular college board programs had declined, AP volumes rose.

It was not until 1961 that we considered English as a Foreign Language as a language for testing. The test of English as a Foreign Language (TOEFL) had its genesis right here when Georgetown hosted the 'Conference on Testing the English Proficiency of Foreign Students' in May 1961. The first director of the TOEFL program, David Harris, was also borrowed from Georgetown University. The annual examinee volume for that test each year is approximately half a million, far more than for any other foreign language test.

By 1941, the multiple choice format had been introduced for foreign language tests as for all college entrance examination board tests. The Spanish test of 1941 had 160 multiple choice questions.

Part I (50 questions): Five Spanish words or expressions are given at the right of each of the following English words. Select the one Spanish word or expression that is the best translation of the English word and blacken out its number on the appropriate line in the answer booklet.

Part II (50 questions): Five English words or expressions are given at the right of each of the following Spanish words or expressions. Select the one English word or expression that is the best translation of the Spanish word or expression and blacken...

Part III (30 questions): Each sentence is followed by five numbered Spanish words or phrases. Select the one Spanish expression that is the best translation of the English word or words in parentheses and blacken... Example:

Entraron las tres señoritas (without our seeing them): 1. Sin nuestra viéndolas 2. Sin verlas 3. Sin que las vean 4. Sin que las veamos 5. Sin que las vieramos

By 1954 the language tests had changed somewhat from the 1941 format, but not much. Vocabulary still ruled and was tested via translation, foreign language to English:

zweig: 1. silence 2. discord 3. branch 4. twin 5. doubt

and English to foreign language:

already: 1. alors 2. déjà 3. surtout 4. encore 5. presque

There were also items that required definitions, others that presented a sentence in English with four suggested translations. Even in 1954, there was

concern about 'other' things that should be tested. The descriptive booklet stated:

Many schools and colleges wish to have their language students acquire much more than reading proficiency. Often they want their students to learn to speak and write the language and to understand the spoken language. Sometimes they expect students to have some familiarity with the literature written in the language and an appreciation of the history and culture of the people who speak it. The college board recognizes the great value of these aspects of language study and could test them, but not under the conditions required by the college entrance examination program.

Great changes were soon to come. Three years later, Sputnik was launched. The impact on language teaching and on the technology applied to language teaching soon was felt in measurement as well.

In 1919, the college entrance examination board received requests from a number of member schools asking that tests of aural comprehension be included in the modern language testing program. In 1921, a committee was formed to investigate the feasibility of aural examinations in the modern languages. The college board, in a frenzy of activity, investigated. In 1962, listening comprehension tests were introduced in the supplementary testing program. By 1971 (a short half-century later), a listening test had become part of the modern language tests in the regular program (except for Hebrew). The listening tests remained a part of the program for only a few years.

At about the same time that the college board was considering the introduction of a listening test to respond to pressure from the language teaching profession, other testing programs were providing for the evaluation of both receptive and productive skills. In the 1960s, the United States Office of Education awarded a substantial grant to the Modern Language Association of America for the development of batteries of tests in French, German, Italian, Russian, and Spanish. The MLA then contracted with the Educational Testing Service for actual test development and operation of a testing program.

I was teaching in New Providence High School in New Jersey when the MLA tests were undergoing trial. Everyone was eager to participate in the norming administrations of the tests. We wanted tests that would reflect what and how we taught. The MLA program offered two levels of tests, lower level for one or two years of high school or one or two semesters of college study and an upper level for three years or three semesters or more of study. In each language there were two parallel versions of each battery with tests of listening comprehension, reading, writing, and speaking. The MLA program also offered test batteries for 'teachers and advance students'. These batteries included tests of culture and applied linguistics in addition to the skills tests. These were the tests used at entry to and exit from the NDEA Foreign Language Institutes. Shortly after the introduction of the MLA tests, a similar testing program was developed for school use by Paul Pimsleur.

The ACTFL proficiency guidelines and the oral proficiency interview have been much debated in recent years. A number of researchers and psychome-

tricians have been concerned about the reliability and practicality of the interview and have recommended the use of 'semidirect' speaking tests. In a way, the circle has been closed. The old MLA speaking test were 'semi-direct' tests. Indeed, some of the item formats in such recent 'semidirect' speaking tests as the test of spoken English are identical to those in the MLA. The scoring scheme and scale for the MLA speaking tests were fashioned on the model of the Foreign Service interview. Senior FSI faculty served as consultants to the MLA. The ACTFL guidelines, scales, and oral interview grew out of the Interagency Language Roundtable (ILR) proficiency level descriptions and interview procedure that, in turn, had evolved from the original FSI scales.

That the MLA tests and, belatedly, the college board tests provided for evaluation of oral skills simply reflected the emphasis on those skills in the classroom. It was not lack of appropriate technology that stopped us from testing oral skills, it was lack of will. The gramophone had already been around for some years when a test or 'aural comprehension' was requested of the college board in 1919. A human voice with a script had been around for a few millenia before that. The Army Language School, now the Defense Language Institute, used 78 RPM records as an aid in teaching uncommon languages during the Second World War. The test of English as a Foreign Language sent 12-inch discs around the world for the listening portion of the test until audio tape, and later audio cassettes, became readily available in the 1970s.

Current technology is impacting measurement as it reflects the use of technology in instruction, e.g. the computer. Just as the computer allows students to progress at their own pace in learning, so too does it allow examinees to escape a lockstep test that requires them to work on masses of material too easy or too difficult to provide much useful information on their performance. Recent work on computerized adaptive testing and computerized mastery testing suggests that computer adaptive tests will be more efficient in time and resources than paper-and-pencil tests, and more 'examinee friendly'. Following an experimental administration of a college board computerized placement test, students were asked to react. Ninety-eight percent found the use of the computer 'very easy' or 'reasonably easy'. Ninety-nine percent reported that the directions given by the computer were 'very clear' or 'reasonably clear'. Forty-eight percent believed the computerized tests to be more challenging than paper-and-pencil tests, thirteen percent said they were less challenging. Sixty percent thought the computerized tests 'less tiring' and only eight percent 'more tiring'. Seventy-four indicated that they would prefer taking a computerized tests over a paper-and-pencil test. Ten percent held the contrary view. An important factor, no doubt, was the length of the test. The computerized test is significantly shorter than the traditional test.

Promising work on computerized foreign language tests is going on at various government agencies. To me, the most exciting possibilities lie in the combination of video disc with computer for testing language skills in a real-life context, together with all the possibilities of random access.

At ETS, we developed an experimental English foreign language test using video disc and computer. Interesting. Expensive. Perhaps the most expensive

piece of the test was the cost of creating and ‘shooting’ a script. Perhaps that is not the most fruitful direction in which to proceed. For a full day, ETS looked like a back lot in Hollywood: camera crew, actors, sound engineers, directors, microphone booms, monitors, clappers. Dollars. There is real life on miles of video tape being recorded constantly all around the world. That may be the best source of material for testing as well as for instruction.

Technology, we love it. If a machine can do, let it. If there is no machine to do it, invent one. We use machines to open cans and sharpen pencils. Then we invent other machines to exercise the muscles that atrophy because they lack the workout they got sharpening pencils and opening cans. We also expect more from our machines than they can possibly deliver.

The following is from a basic text on the use of audiovisual aids in the classroom. It was published in 1957. It describes the dual-channel tape recorder used in the language laboratory at the Georgetown University Institute of Language and Linguistics.

The principle behind it is simple. The tape is divided into two channels by the use of twin magnetic heads. With the dual channel system, the instructor’s voice can be recorded and heard on one channel. The same operation can be performed by the student on the second channel. Through the process of connecting the first earphone of the headset to one channel, and the second to the other, the student can listen at will to both channels alternately or simultaneously. The instruct can record his exercises on one channel, and the student his emulation on the second. Through a process of comparison and repetition the student in time achieves by ear a perfect pronunciation of the instructor’s sentences. This is the basic part of the method.

Today we have ‘interactive video’, the omnipotent computer wed to the vampish video disc. Irresistible.

Last year I worked with the faculty at the United States Naval Academy in the development of IAV lessons for their Spanish program. These lessons I learned about IAV:

1. It is expensive to produce lessons.
2. It takes a lot of time to develop a lesson.
3. It requires much creativity and imagination.
4. IAV lessons can be deadly boring.
5. IAV lessons can be exciting and effective.

The faculty involved in creating IAV lessons, despite their great enthusiasm, found themselves laboring under a greatly increased workload combined with tight production deadlines; all the while enjoying the full-time dedication to the task of team leader, and a programmer who is also a linguist.

The faculty were all excellent language teachers, creative and energetic. That did not guarantee that all our lessons were creative, exciting, or even very interesting. It struck me that maybe what is needed is a collaboration between us, the language specialists, and the video whiz kids. I was ashamed of the

pedestrian work I did with the medium, especially at the beginning. I know that I did not exploit its potential even minimally.

If there is a recurring complaint about technology, it is the lack of imagination in its use. What we too often see is a workbook on a screen. We need help in finding the answers. We need partners.

I will turn to my greatest concern: the kids in French-I in Hopewell Valley Central High School or Italian-II at Trenton High, or Spanish-II at Notre Dame High School, just down the road from me, and their teachers. I requested of the ERIC Clearinghouse on Languages and Linguistics a couple of computer searches, 'incorporating CAI into the second language curriculum' and 'teaching second language with computers'. In each there were hundreds of entries. ERIC, with its usual efficiency, had the searches on my desk days after a phone call. A couple of days later, I was speaking to about 200 language teachers. I asked how many were acquainted with ERIC and its services. Two people responded. One percent. I do not know if that is typical across the country. I do not know if it was an aberration. I do know that it is a shame. In the world of *GURT*, ERIC is an old pal. But in the world of three sections of *Español*-I or II, lunch duty, and supervised study hall, ERIC is a stranger. So too are CAI and CAT and IAV.

I do not presume to have the answers as to where to apply limited resources. I know that it was a joy to teach my classes of a dozen midshipmen, who, after class, could interact with the computer and the video on lessons my colleagues and I had made. It was a thrill. It is also an experience that my former friends and colleagues at State Street Junior High School will never have, at least not in my lifetime, I am sure.

This is from Al Shanker's column, where he was criticizing the Whittle communications proposal to provide free TVs and VCRs to schools in exchange for air time.

Many schools around the country would love to get their hands on some modern educational technology, but they just can't afford it. I'm not talking only about expensive and sophisticated computers. Most classrooms don't even have a VCR or TV set as permanent classroom equipment. Of course, the school usually has such items, but there are often scheduling problems, and equipment that is moved from floor to floor and room to room is often out of commission.

That, I am afraid, is reality. It is very easy to be discouraged. I just received an 'invitation to participate and/or call for papers' for an International Conference on Visions of Higher Education: Transnational Dialogue: Higher Education and the Communication Environment. The conference will be held in Lochau, Austria. Among the topics to be focused upon are:

Machine and humans: An unequal starting point
Mystification of technological progress

Plane fares, lodging, refreshments, work hours, will represent a significant investment, I am sure. But it is in the interest of higher education. I would like to see a bit more investment in 'lower education', investment that might get a few more VCRs and TVs into the classrooms of Detroit and Newark and Camden. It is discouraging to know that while scholars are discussing the 'Mystification of technological progress', I will pass padlocked doors and an armed guard to enter a nearby school to visit a language class where a set of audio cassettes is unaffordable.

Now, after sounding my gloom and distress, let me put on a happy face. I picked up the proceedings of the *Georgetown University Round Table on Languages and Linguistics 1973*. One of the interest groups sessions was on 'Coordination of high school and college objectives in foreign language instruction'. Much discussion centered around the need for commonly accepted descriptions of levels of competence. The Foreign Service Institute scales were mentioned as an exemplar. Robert Lado, who was in his last year as Dean of this school, commented on the FSI scale. He said that:

- (1) It is difficult to discriminate well at the lower levels since many different degrees of performance are represented within a single rating.
- (2) The profession need not agree on a test but it should agree on the definition of standards.
- (3) It is time to set down standards, define levels, and then find the means by which we can measure them.

Sixteen years have passed. We may not all agree with everything that has taken place under the umbrella of 'proficiency'. I think we must agree though, that what was said here in 1973 was heard, and helped to refocus national interest in foreign language education and to return excitement and a sense of mission to our profession.

Challenging teachers and harnessing technology

June K. Phillips

Tennessee Foreign Language Institute

This paper does not deal in hardware or software per se, nor does it expand the acronyms by which we now communicate, be it CALL, CD-ROM, IAV, CAT, SCOLA, CALICO, PICS, or TWIT; it introduces no new women to the ranks of ATHENA, CALLIOPE, ELIZA, or MONIQUE, no image-evoking names like VISTA or VELVET or EIDS (spelled with an *E*).¹ Roundtables, conferences, articles that portray the progress and programs on the leading edge which are being developed by the new technocrats allow us to dream of what foreign language teaching might be. But out there, away from this rarefied atmosphere, in places such as Bell Buckle or Dismal, Tennessee, the gap between technology offerings and daily teaching practice widens with the passing days. This article presents a challenge to harness technology to serve those teachers and their students with pedagogically sound and linguistically appropriate materials.²

The quantum leaps in the types of technology, the variety of systems and the range of materials available fail to be reflected yet in the routine practices of foreign language teachers. Indeed, if we are to learn lessons from the past and not be doomed to repeat its mistakes, then we must with all haste transmit, translate, transfer into the hands of practitioners not just the outward glitz of these new technologies, but also the sound pedagogical assumptions that the best of these exemplify. The future requires that foreign language teachers change what they do and adapt rapidly to the offerings of technology or they will be left behind as aspiring language learners look elsewhere for their instruction.

In many instances, the corporate and business world have already abandoned traditional instruction as delivered by educational institutions. Readers of airline magazines discern that the appeal to language learning which they pitch to adults, and to their children, is predicated upon combinations of video, audio, and novel methods aimed at communicative goals. Astute corporate learners do not really believe that these programs prepare them to speak like a diplomat in thirty days, but they do buy into ones which blend technology and teaching in ways which accommodate their time constraints and their reasons for learning. The language teaching profession might benefit from observing and evaluating these systems rather than dismissing them as commercial intrusions into the educational arena.

To prevent the formation of an unbridgeable chasm, an agenda for bridging the present gap between technology and foreign language teaching must be enacted. Technology changes rapidly; teaching and teachers change slowly. Even in terms of basic instructional design, today's schools cover the spectrum; for each communicatively oriented classroom where a teacher

makes decisions based upon relevant theories and practices of second language acquisition, many more classrooms consist primarily of mechanical drill devoid of context, lockstep learning, and little use of authentic listening or reading materials. These latter often pride themselves on students who respond in full sentences about unknown, uninteresting, contrived characters named Pierre, Pieter, or Pedro, and who score well on the competitive grammar tests promoted by language organizations in the states. When 'technology' appears, it comes in the guise of tapes which accompany the text (in 'repeat in the pause' format), an occasional film (and believe it or not the filmstrip still lives) or more frequently now, a video (illegally copied from a colleague) of the *Karate Kid* in Spanish. (An amusing curiosity? Yes. Culturally appropriate or justifiable? Probably not.)

Why this chasm? What can be done to harness technology and to challenge teachers to use it wisely in promotion of their goals? Most importantly, can pedagogy serve as a linkage to assure that the wonders of technology take account of existing research in second language acquisition even as they in turn contribute to that research base? Can teachers be taught to use technology more effectively?

Challenging teachers. On the very positive side, teachers are fascinated by the possibilities for the enhanced foreign language environment that video, computers, and the combinations thereof provide, and they aspire to use them in their teaching. Part of the difficulty in implementation lies in the fact that teachers are being exposed to technologies from which they have never learned themselves. The fact remains that they 'overcame' language in spite of outdated methods, contrived materials, and an intensive emphasis on form over function in an environment consisting of teacher and text, paper and pencil, blackboard and chalk. While their head may acknowledge the logic of more communicative approaches and the insights gathered by research on acquisition, their hearts have not been fully convinced that what worked for them might not work for others. Changing teacher behaviors, especially in the use of technology, requires more than one-half-day computer workshops or a session on video in teaching at a conference. Research, especially that of the ethnographic type, demonstrates that (1) the greatest influence on teaching is the instructor's personal history of learning (Britzman 1986), and (2) teachers possess underlying theories of teaching which exert major influence their practices. Furthermore, those positions may be incompatible with practices they want to or claim to espouse (Swanson-Owens 1986).

Equally challenging for teachers is that technologies and the shrinking world have brought spontaneous and unedited target language to their doorsteps. Teachers' own foreign language skills are compared, measured, and evaluated because native speech and authentic materials are available at the push of a button or on the magazine shelves of neighborhood bookstores. News broadcasts from around the world require that professors and teachers deal with content areas outside the traditional scholarly fields of linguistics or literature. Often they are prepared neither with the language nor the social sciences to understand and to explicate for their students live reports on Central American politics or chemical pollution of the Rhine. Teachers now confront having to listen to or read texts for which they cannot define every

word, and they do not have time to look up either words or background information without violating the timeliness factor which renders the content and language interesting and important.

An additional drawback to the most effective integration of technology and the newer visual, audio, and print materials on the market lies in an assumption that these items serve primarily peripheral, supplemental, remedial or enrichment functions. They can perform that function, but they are equally capable of replacing or substituting for more routine or mundane classroom activities. The lament that 'I want to use more video or magazine readings, but I just don't have time . . .' can only be eradicated by an understanding that watching a TV segment or reading a newspaper item builds skills and enhances language acquisition to the degree that one need not also cover a pabulum-like recombination narrative from the text.

The essence is that foreign language educators, in consort with technologists, must provide the kinds of preservice and inservice opportunities that develop understandings of the underlying theories and assumptions in new systems and not just the tricks of using them. Unfortunately, observations in the schools and extensive contact with teachers reveal the current pattern to be one in which the materials actually used in many of today's classrooms reflect the least sophisticated methodological underpinnings and the least effective integration of technology and curriculum. While that is bound to occur when the most advanced technologies remain in an experimental and developmental stage, the reality is that as long as the 'drill and kill' software, the vocabulary games, and the electronic workbook dominate computer use in the foreign language classroom, little impetus to go beyond that will occur. These mechanical programs will become as institutionalized as the pattern drill.

An equal danger arises as authentic audio and video programs appear for which a complete script is always provided so that neither student nor teacher has to rely on the sound or scene. Used well, scripts facilitate the teacher's preparation of listening materials; abused, they mask the teaching of the listening process and create impossible expectations for teacher and learner alike. Achieving balance is the goal, thus listening tasks based upon what is heard, not upon what is read, must be in the instructional program. These richly contextualized and authentic materials ideally provide practice in the process of comprehending; yet accompanying support materials, too often ill designed, focus on vocabulary lists (English/target language) or elicit discrete point grammatical forms as responses to questions about language. Materials used in this manner sustain an approach based upon language forms amid all the lip-service to communication, proficiency, and cultural awareness.

The computer's unfulfilled promise. Computer use in the average foreign language classroom remains minimal. Recent surveys such as one conducted by *USA Today* reveal that although many of the nation's students now have computer literacy courses, learners use the computer infrequently in any discipline. Software reviewers testify to the range of computer activities that can be pursued. Wyatt (1987) credits CALL with being able to perform 'Instructional', 'Collaborative', and 'Facilitative' tasks with students. The instructional mode includes the tutorial, practice drill, and games most

commonly seen in classrooms. The collaborative involves higher level tasks of modeling, discovery, simulation, adventure reading which encourages students to interact with the computer. The facilitative incorporates functions such as word processing, spell check, and thesaurus which have no direct pedagogical application.

Becker (1989) also lists the common software types available to foreign language teachers as (1) drill and practice, (2) games, (3) tutorials, (4) simulations, and (5) problem solving. The first three types are all language-centered, usually require totally correct response, and they still predominate the field. The latter reach into higher level skills and focus more on meaning and communication, but few students work with this level of software in their foreign language courses. An informal survey of the software review columns now prevalent in most journals and newsletters discloses that, in spite of all the possibilities, the realities favor programs with a mechanistic focus. The longer we tolerate computer usage that remains in the realm of entertainment or remediation, the wider the gap between the best and the worst utilization of this technology.

What about the argument that the computer removes tedious drill from the classroom, thereby freeing teachers and students for communicative activities and more creative language use? An interesting argument, but if stimulus-response with language patterns and detailed error correction are not the most effective means to promote communication in the classroom, why should they be more satisfying in computer formats? Is it simply that old practices are hard to abandon, that the role of context and function is not yet firmly established? Is there not an inconsistency in talking about tolerance for error and interlanguage on one hand, then sending the kid to a computer program which accepts only one well-formed answer? (Caveat: This does not refer to the humane, tolerant, cue-giving programs, only to the 'Drillmasters' which seem to dominate the field.)

All is not dismal, and computer-wise foreign language teachers create new programs regularly. Likewise, pedagogically aware teachers do not relegate unsound, unfounded, and outdated practices to the computer; instead, they hunger for and anxiously await the transfer of the innovative programs about which they read and hear to the realm of the affordable. In the meantime, teachers must be challenged to demonstrate their commitment to communicative language teaching by using technology to enhance that goal.

Ariew (1987) provides a cogent and well explicated summary of the advantages and disadvantages of computer and video in the proficiency-oriented classroom. These systems can be utilized to enrich a whole language environment. To do so, the methodology of teachers must have first undergone change; only then will it be capable of integrating the best of what technology has to offer.

Video and the TV generation. Many view video as the most promising technology on the horizon; little about it is foreign to either students or teachers. It introduces the sound and sights of language in authentic contexts and for real-world purposes. New sources for video materials at reasonable cost appear regularly, and many of these materials are accompanied by instructional guides. Material developers, educators, and researchers

interested in listening comprehension have joined forces so that the pedagogy in the supporting materials reflects insights from studies in receptive skill processing. Altman's (1989) book, *The Video Connection*, provides an excellent theory-to-practice resource.

The challenge with effective use of authentic video and audio lies not in the carefully developed programs themselves but in total reliance upon them. Foreign language instructors must exploit video as more than an opportunity for controlled audiovisual text and develop listening skills in extensive as well as intensive ways parallel to liberated reading formats. The demand for scripts to accompany video and audio materials should be accommodated discretely. Balance should be sought so that teachers understand that some telecasts might be preserved and used for detailed listening, but for the majority comprehension is dependent upon 'hot' topics that allow students to develop strategies for instant viewing with an expectation of understanding as much as possible and a realization that it will not be every utterance.

Supporting materials should go beyond scripting to processing strategies which develop understandings of how learners listen, of how they can be led to listen, and why and when partial understanding is perfectly acceptable. Ideal formats would involve integrating listening materials with theoretical premises such as those posited by Byrnes (1984) and Long (1989) and with pedagogical principles such as those proposed by Glisan (1988), Joiner (1986), Joiner, Adkins, and Eykyn (1989), and Weissenreider (1987).

Research with technology. We are entering a stage where technology and supporting materials are being developed with a basis in relevant pedagogical principles, and the technology itself is being exploited to teach us more about language learning. That is an ideal exchange rich in potential new information. Studies such as Robinson's (1989), which used the computer to test out different instructional treatments and to ascertain the effectiveness of various error feedback modes, represent the type of research possible with CALL programs especially. Progress can be made by evaluating and then removing from the market ineffective software. Teachers need data on how learners interact with programs. It is not enough that students are quiet or noncomplaining. We must stop banking on their intrinsic fascination with the machine to do mindless things. The next generation will not tolerate electronic workbooks as the present one does.

The advent of more interactive video programs, facilitated by software such as Hypercard, enables computer-wise teachers to construct their own materials. They must do so with caution and with attention to research on the effectiveness of their product. Worst case scenarios include demonstrations of a hypercard-generated program to teach fruits and vegetables in Spanish. The result of intricate technology failed to surpass that of an illustration in a workbook; intellectually, a sophisticated computer program became the stimulus for a 'labeling activity' by students.

Best cases, such as that planned by the developers of ATHENA (Furstenberg 1988), submit the interactive program to research on student attitude, satisfaction, fatigue, motivation, in addition to learning activities that reach into decision making and problem solving. Given the huge costs associated with IAV, it would be disastrous to discover that a few hours of

student/technology interaction intrigue, additional ones bore, and the full program turns learners off or they tune out. The research needs to be done in the development stage.

Reports of careful research and experimentation with technology as a learning tool are increasing. Garza's (NFLC 1988) work with 'reverse captioning' constitutes an example of focusing on a technology, in this case video, and developing a device to provide assistance to comprehension. His reverse captioning consists of adding native-language subtitles to an authentic video segment; he then plans to assess the effect of this comprehension cue. More projects of this scope and focus would be welcome contributions to effective use of technology.

In sum, the high costs associated with IAV and other technologies demand that research on effectiveness be conducted *a priori* and not *a posteriori*. Certainly, one painful lesson from the past is still visible in the abandoned language laboratories whose booths stand all in a row gathering dust and calling out to administrators unwillingly to refurbish them for yet another great experiment.

An agenda for teachers and technology. The time to set an agenda, and more critically to act upon it, is now if we are to bring technology and teaching into harmony for the benefit of language learning. That agenda might include:

- Developing technologies which instill greater communicative and content competencies in teachers themselves as learners of other languages;
- Training teachers with the products and processes of technology so that they transfer that teaching history to their instructional approaches;
- Making wiser decisions on the best use of software by determining effectiveness at promoting higher cognitive levels of language learning;
- Bringing the best of pedagogy to serve technology and vice versa; creating partnerships of designers, teachers, teacher educators;
- Sharing efforts so as not to reinvent, to duplicate or otherwise feel a compulsion to do everything; spread the expertise;
- Finding funding sources to absorb some of the costs;
- Fighting 'future-phobia' in the profession; that means immersing people, not drowning them.

In the past, the profession has often moved at a snail's pace either out of fear of change or extreme caution in initiating action because of the lack of a strong theoretical base. Today, education may need to replicate the sciences more closely and proceed simultaneously to experiment on both fronts. Pursue the theoretical models without delaying exploration on the level of practice. Otherwise, technology will outdistance our ability to harness it for meaningful use. We need to design, to experiment, to evaluate, based on the best thinking and theories of the moment.

The world beyond educational institutions is venturing into new ways of using technology to teach languages, to translate, to communicate in an international forum. As language educators, as linguists, as multilingual

individuals, we should be connecting our knowledge with theirs and capitalizing on shared visions for our mission.

Notes

1. The acronyms and programs cited refer to: CALL (Computer Assisted Language Learning); CD-ROM (Compact Disc, Read Only Memory); IAV (Interactive Video); CAT (Computer Adaptive Testing); SCOLA (Satellite Communications for Learning Association); CALICO (Computer Assisted Language Learning & Instruction Consortium); PICS (Project for International Communication Studies); TWIT (Two-Way Instructional Television); ATHENA Language Learning Project at MIT; CALLIOPE (Computer-Assisted Language Learning Outreach Project for Education); ELIZA (MIT program named for the woman in Shaw's *Pygmalion* who was taught to talk right); MONIQUE, French Drill and Practice Driver program; VISTA, Word Processing Program in Spanish to develop writing skills at South Manchester Community College (UK); VELVET (Video Enhanced Learning, Video Enhanced Teaching); EIDS (Electronic Information Delivery System).

2. Articles typical of these include: Joel Schwarz, 'Speaking your language' in *USAIR*, May 1988, 90-97, on machine translation for international business people, a program developed at Carnegie-Mellon University's Center for Machine Translation; Brad Buchholz, 'Sisters of other tongues' in *American Way* (American Airlines), 30-38, an article about a Japanese language course for Motorola employees developed by International Language Services, using the Lozanov method.

References

- Altman, Rick. 1989. The video connection. Boston: Houghton Mifflin.
- Ariew, Robert. 1987. Integrating video and CALL in the curriculum: The role of the ACTFL Guidelines. In: Modern media in foreign language education: Theory and implementation. Ed. Wm. Flint Smith. Lincolnwood, Ill.: National Textbook Company. 41-66.
- Becker, James E. 1989. The computer and foreign languages in the high school. In: Modern technology in foreign language education: Applications and projects. Ed. Wm. Flint Smith. Lincolnwood, Ill.: National Textbook Company. 135-59.
- Britzman, Deborah P. 1986. Cultural myths in the making of a teacher: Biography and social structure. *Harvard Educational Review* 56.442-56.
- Byrnes, Heidi. 1984. The role of listening comprehension: A theoretical base. *Foreign Language Annals* 17.317-29.
- Furstenberg, Gilberte. 1988. Quand une pedagogie rencontre une technologie... Le français dans le monde (aout/septembre) 158-66.
- Glisan, Eileen W. 1988. A plan for teaching listening comprehension: Adaptation of an instructional reading model. *Foreign Language Annals* 21.9-16.
- Joiner, Elizabeth G. 1986. Listening in the foreign language. In: Listening, reading, writing: Analysis and application. Ed. Barbara H. Wing. Middlebury, Vt.: Northeast Conference on the Teaching of Foreign Languages. 43-70.
- Joiner, Elizabeth G., Polly B. Adkins, and Lollie B. Eykyn. 1989. Skimming and scanning with Champs-Elysées: Using authentic materials to improve foreign language listening. *The French Review* 62.427-35.
- Long, Donna Reseigh. 1989. Second language listening comprehension: A schemata-theoretic perspective. *Modern Language Journal* 73.32-40.
- National Foreign Language Center. 1988. Research in progress. Using authentic video in the classroom. NFLC perspectives.

- Robinson, Gail L. 1989. The CLCCS CALL study: Methods, error feedback, attitudes and achievement. In: *Modern technology in foreign language education: Applications and projects*. Ed. Wm. Flint Smith. Lincolnwood, Ill.: National Textbook Company. 119-159.
- Swanson-Owens, Deborah. 1986. Identifying natural sources of resistance: A case study of implementing writing across the curriculum. *Research in the Teaching of English* 20.69-96.
- Weissenreider, Maureen. 1987. Listening to the news in Spanish. *Modern Language Journal* 71.18-27.
- Wyatt, David H. 1987. Applying pedagogical principles to CALL courseware development. In: *Modern media in foreign language education: Theory and implementation*. Ed. Wm. Flint Smith. Lincolnwood, Ill.: National Textbook Company. 85-98.

Building on the past: New directions in CAI/IL

Frank Otto

Brigham Young University

Introduction. It is a privilege to be here and participate in the 200th anniversary program honoring the establishment of Georgetown University. The title of my presentation is indeed rather ambitious but there is a definite need to look at the teaching and learning of languages and the many ways in which technology can enhance options available to us and our students.

There are certain principles that underlie effective use of computer-assisted instruction (CAI), especially as we apply high technology to provide a truly interactive and individualized learning environment. We must emphasize the fact that technology is not intended to replace the master teacher but is designed to assist us and our students by providing a variety of realistic language learning experiences. The teacher certainly is the director of the learning environment even though students gain a good deal of learner control of options consistent with individual learning styles. Learning is discovery oriented and based on the concept of mastery. Technology continues to offer us more and more learner options. We continue to see increased teacher involvement, research, and interest in the status of technology and the role that it can play to enhance the learning environment.

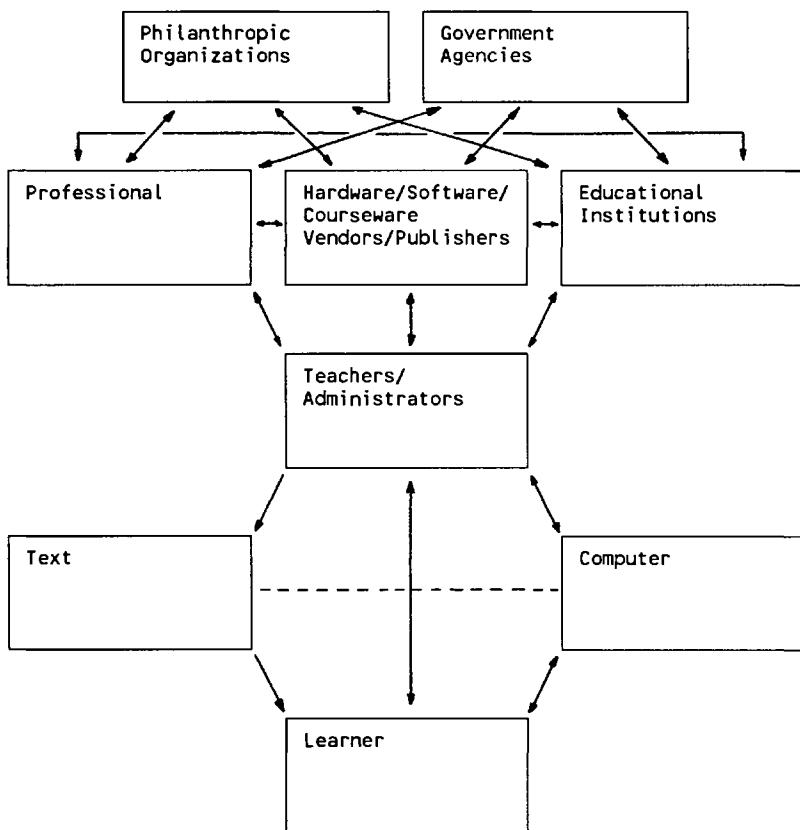
As the result of over 15 years of experience with research, development and instruction dealing with trends, issues and capabilities of CAI systems, I have learned some lessons of great importance: Competent teachers and administrators must be centrally involved in designing and managing the teaching/learning process, and a major trend in teaching and learning during the past 30 years has been toward the individualization of instruction. This quest has not only altered the teaching and learning environment but has directed the attention of those interested in CAI to a new area of emphasis most commonly referred to as interactive learning. Thus, we have yet another acronym to add to a long list of labels that have appeared in our professional literature during these past four decades: CAI/IL (Computer-Assisted Instruction/Interactive Learning).

The new endeavor calls upon all of us to work cooperatively as members of a team dedicated to providing the most challenging and rewarding interactive learning environment. (See Figure 1, which displays the personnel and organizations involved in CAI/IL environments.)

Notes to accompany Figure 1, the chart which displays the personnel and organizations involved in CAI/IL environments. For those who have been involved in courseware development projects involving applications of high

technology, it becomes more and more apparent that the resulting materials must come about with the participation and cooperation of many individuals involved in designing, developing, programming, and evaluating packages for use by resident and nonresident students.

Figure 1. Personnel and organizations involved in CAI/IL* environments.



* Computer-Assisted Instruction/Interactive Learning Environment

Because the development process must be very carefully planned and pursued with great attention to detail and quality control, and because the personnel involved encompass many areas of expertise, it is recommended that those developing CAI/IL packages work with the support of outside funding agencies unless a large consortium of institutions can be involved in order to share the costs of an exemplary and comprehensive courseware development project.

It has become increasingly more important for professional language organizations, hardware/software courseware vendors/publishers, and

educational institutions to work closely together in order to develop programs that will be of maximum benefit to teachers/administrators utilizing courseware packages. While professional language organizations have tended to play a rather passive role regarding text/courseware development in the past, they must now become much more assertive in assisting interested professionals in realizing their goals for the development and use of pedagogically sound and cost-feasible courseware programs. This requires the close cooperation (and in many cases direct assistance) on the part of the hardware/software/courseware vendors and publishers. It becomes increasingly more important for educational institutions to reduce the gap between potential 'real world' learning situations and the training programs developed for prospective as well as in-service teachers. It is only with the close cooperation of professional language organizations, hardware/software /courseware vendors and publishers, and educational institutions that we can hope to meet the real needs of teachers and administrators utilizing technology as an integral component of the teaching and learning environment.

We all know about the importance of hardware, the software that drives it and the courseware with which students learn and instructors teach. We are sometimes frustrated by the vaporware that is either delayed or delivered in a format that is not totally acceptable. But the most significant '-ware' of all is *warmware*--the teachers and students involved with CAI/IL without which none of the other '-wares' would matter.

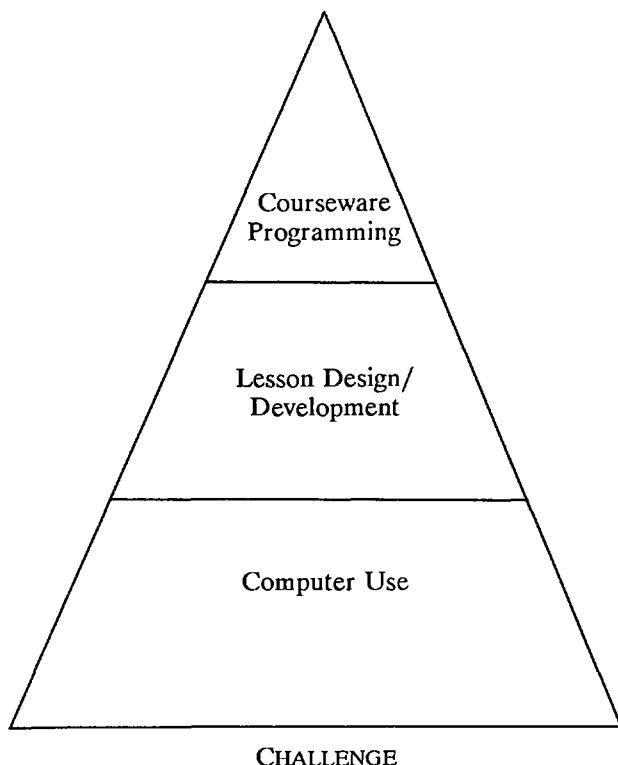
Challenge. From the beginning, it must be admitted that several experts question whether teachers should be expected to program courseware. They contend that the teachers' expertise is so valuable that it should not be wasted by asking the teachers to also be programmers. Certainly, teachers' experience interacting with students and with the subject matter is essential if good courseware is to be produced. We can, however, overemphasize the value of teachers' previous training and experience, for it is usually with traditional classroom experiences instead of with varied CAI media. Learning the techniques of interaction appropriate for CAI is a significant area of education that teachers will have to acquire if they are to prepare to develop CAI courseware from scratch. The use of Courseware Authoring Template Systems (CATS), however, makes it possible for teachers to develop full-featured CAI courseware without having to understand fully CAI media capabilities, since the essential information is either built into the system or explained and prompted for in the design forms and information sheets from which they work. (See Figure 2.)

The programming phase of courseware development is likewise simplified to the point that it becomes feasible for teachers to accomplish these tasks by using the Author program of CATS.

When using CATS authoring environment, the first decision to be made is 'What kind of courseware is to be developed?' Ultimately, when a selection of CATS is available, the courseware developers will select the one that was designed for the type of courseware they plan to develop. With the selected CATS in hand, the developers then use their expertise in the language and in classroom interaction to fill in the courseware design forms. With the design

forms complete, all that is necessary to complete the lessons is available, the programming is already complete, and only the lesson-specific information needs to be added. Once programmed, the lessons run bug free, since the bugs have already been worked out on prototype lessons programmed with the CATS.

Figure 2. Levels of high-tech involvement.



CAI/IL programs.

- (1) CAI/IL programs should be used which allow for individualized and personalized student interaction not possible in a normal classroom setting:
 - Through student control of rate of presentation
 - Through availability of help features
 - Through individualized branching for remedial instruction
 - Through the program's ability to adjust the level of difficulty according to the student's proficiency
 - Through student ability to continue receiving instruction and practice until a level of mastery is attained
 - Through immediate correction of errors so that mistakes will not be practiced and become habituated.

(2) CAI/IL programs can provide for language-learning experiences not otherwise available (such as actual use of the language in a foreign country) through simulations provided through the 'subjective camera approach.'

(3) CAI/IL programs with a record-keeping function provide both the student and teacher with ongoing evaluation that would be impossible in the conventional classroom setting.

(4) CAI/IL programs provide a link with the real world through simulated, culturally authentic events and situations in which students role play and react in an interactive manner by taking advantage of branched alternate paths as well as a variety of learner options within each path.

Guidelines of effective CAI/IL programs. While there are many guidelines that may be listed for effective CAI/IL programs, the following are the most important.

(1) *Perceived purpose.* Learners learn best if they can see or develop for themselves a reason for learning the material being presented.

(2) *Appropriate practice.* Quite simply, 'appropriate practice' means doing what the objective calls for. If our objective is that the learner will be able to obtain services in a restaurant (order food and pay for it), then the best possible way to achieve that objective is for the learner to practice in a role-playing situation.

(3) *Graduated sequence.* 'Graduated sequence' involves progressing from the easy to the difficult, from the familiar to the unfamiliar, from the simple to the complex, from the clearly stated to the implied. It is based on the assumption that success is a very important factor in the learning process.

(4) *Knowledge of results.* While we learn by doing, we can learn to do things incorrectly by practicing them incorrectly. 'Knowledge of results' involves telling learners whether they are doing something correctly or incorrectly and, if they are incorrect, telling them as precisely and as subtly as possible what is wrong and how to correct it. This process allows students to avoid habituated errors.

(5) *Individual differentiation.* This principle is based on the assumption that not all learners are equal. They have different entry skills, aptitudes and attitudes. Every effort should be made to permit each learner to proceed at a pace that challenges yet permits success. There should be no correlation attempted between the number of attempts made and the evaluation of student success. This approach effectively reduces the correlation between aptitude and achievement to zero (0).

(6) *Fail-safe protected learning environment.* Students are not permitted to fail. The record keeping function allows them to proceed as long as an acceptable level of mastery is maintained (we use 80 percent in BYU CALI Research projects). When students fall below the agreed upon level of mastery, remediation, review, and additional examples/explanations are provided in order to bring them back on track so that their learning experience will be successful.

Unique features of interactive technologies. In a move beyond the passive linear learning experiences provided by movies and TV, massive amounts of information are processed and made available through electronic means with

a speed, accuracy and efficiency previously impossible to attain. Beyond the processing of data, however, we now have systems for preserving, processing, and accessing audio and video information that are making it possible to share and teach information in ways not available with conventional text materials.

Movies and TV, along with the sophisticated delivery systems such as satellites and fiber optic cables, are making it possible for people to see and hear and thereby experience almost first hand what is happening in virtually any culture, thus allowing students to gain access vicariously to the experiences of other people. Meaning is often conveyed in pictures and sounds, but students do not always recognize it. It is often necessary to provide interpretations and explanations of the visual and audio information before the students learn to recognize meaning fully.

Modern technology is now making information systems available that combine the effectiveness of visuals and sound with the natural learning environment of interaction.

(1) Classroom instruction has always been plagued with the difficulty of communicating to heterogeneous groups of learners. Instruction might be too difficult for slower learners and too slow for rapid learners. Only by individual tutoring could the specific needs of each student be met in the most effective way. Approaching this ideal, modern interactive training systems are designed to respond to each individual student by providing the means for each student to obtain special help on each step in the learning progression. Slower learners will find the many types of help they need, while rapid learners can obtain the bits and pieces needed to move along at their pace. Thus, all students are able to progress efficiently toward heretofore unachievable levels of proficiency.

(2) The communication power of video and audio can be utilized effectively with the newer technologies, each adding greatly to the educational power needed in our modern world. Education is no longer limited to the capabilities of a book. Students are able to learn from many simulated experiences. The computer drives an interactive learning system that can provide the personal guidance needed to make salient those features of the material that the student must come to understand.

These educational possibilities are becoming available because of the development of computer systems and audio/video equipment that were formerly not available.

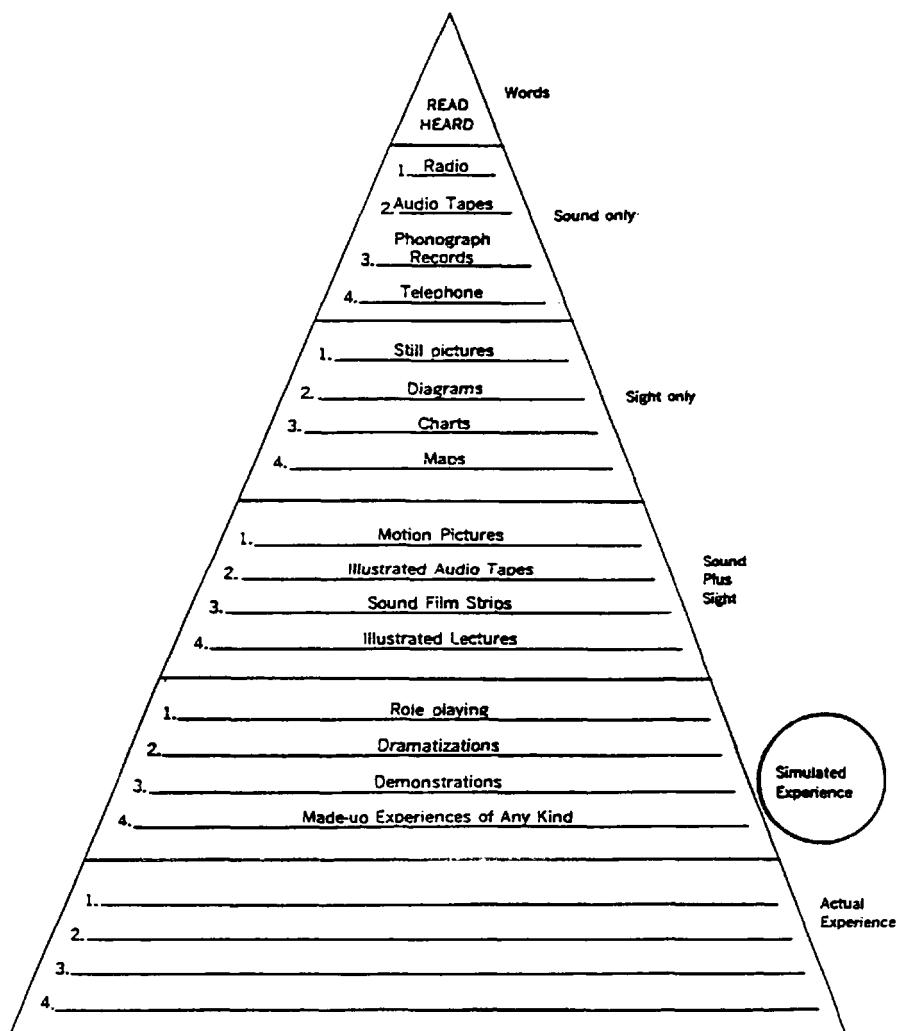
(1) Computer technology has developed so that powerful processors with highly advanced capabilities can be placed on desks or even in our laps.

(2) Videodiscs and videodisc players contain 54,000 pictures that can be either still pictures, slides, or full motion sequence with two tracks of audio. It is possible to show several options available to learners confronting a variety of tasks. It is also possible to present instructional materials in several languages on a single disc.

(3) CD-ROM/CD-I technology offers us a depth and variety of applications that enable us to design and develop the most creative and challenging interactive learning environment imaginable.

Interactive simulation. Effective CAI/IL programs branch to any of several next frames according to the student's response. The next frame may be another frame in the simulation, or it may be a frame providing language information the student needs to master. When the student's responses indicate that the task has been mastered, the program branches to another task in the simulated activity and continues on. Another option is for the program to allow the student to receive feedback from an error by switching into a gaming mode that makes the student aware of the consequences of errors. In either case, each student will receive the training that most closely meets individual training needs.

Figure 3. Cone of learning experiences.



The educational promise of CAI/IL includes the goals of individualizing and personalizing instruction, simulating experiences not otherwise available and enabling students to work at their own pace. When planning instructional specifications, teachers and administrators must choose objectives and select or design courseware that requires student interaction to the fullest extent.

Comments to accompany the cone of experience. According to Edgar Dale, there are six levels of experience extending from words through sound, sight, sound plus sight, simulated experience, and actual experience. In any teaching/learning environment, the experience closest to actual experience is simulated experience; therefore, our efforts to educate learners and to relate that training to on-the-job tasks can best be accomplished by allowing the learner to play out roles that are in keeping with career assignments. This approach is especially beneficial as the essential component of nonresident training programs.

In developing special language-based training programs, simulation materials may best be generated by using guidelines describing tasks that are clearly job specific. The advantages to this approach of working with simulation are economy of time, increased accuracy of specialized language use including key vocabulary, and the fact that surprise/reality can be built into the simulated experience.

Guidelines used to develop an exemplary educational experience provided by interactive video instruction.

- (1) Rapid access to lesson material that is divided and subdivided allows students to work with the portions they need.
- (2) Screen menus list all options for the students so they can quickly select and obtain the types of help needed.
- (3) Rapid access to video material can be repeated as frequently as needed.
- (4) Key words help the student concentrate on the essential matters.
- (5) Helpful information regarding vocabulary, grammar, culture, and idioms are quickly retrieved according to student needs.
- (6) Students can study a transcript of the dialogue while listening to the language or while watching the video material.
- (7) Any part of the video material or audio material can be repeated as often as desired, whether to aid memorization, to aid perception of sounds, or to study pronunciation or meaning or cultural notes.
- (8) Alternate more carefully enunciated audio (MCEA) recordings are available to aid students who cannot yet comprehend the audio at the speed normally spoken by native speakers of the target language.
- (9) Questions are used to direct the students to search for specific information. All of the learning helps that are available elsewhere can be made available to students in their search for answers. Questions are also used in student evaluations to obtain evidence of mastery of the material.
- (10) Foreign characters are based on popular existing video materials as well as prepared footage. The same video footage can be treated in a variety

of ways to meet the needs of a range of students, from beginning language learners to advanced students.

Characteristics of nonresident language training. Language training emphasizing nonresident instruction has two goals: general language proficiency in which particular skills are emphasized, and mission capability in which job-specific skills are stressed. (With the aid of technology, this approach is often referred to as 'high-tech' or 'high torch').

Nonresident language training programs have the following unique characteristics.

(1) *Students must remediate/refresh/maintain/enhance their language proficiency without sacrificing career needs.* This situation is totally unlike a conventional resident classroom orientation in which students leave their 'career' in order to attend an instructional program. Learners must progress in skills to be attained (in this case, increased proficiency in a foreign language) while engaged in full-time employment.

(2) *Nonresident language training programs must be based on job-related linguistic skills and knowledge that is essential to specific tasks.* If the activities in the nonresident program do not relate directly to job tasks, interest will be lost, the program will be viewed as 'boring,' and the experience will not be accepted as one that is either linguistically challenging or professionally rewarding.

(3) *Students are primarily accountable for their progress in a nonresident training program.* Learners in a conventional resident classroom environment rely heavily on the daily guidance and control that the teacher has on instruction. In a nonresident program, the responsibility for daily effort and progress resides almost exclusively with the individual learner. For this reason, every effort must be made to select/develop courseware that provides as much motivation to the learner as possible.

(4) *Instruction in a nonresident learning facility is multilevel.* While students in a conventional resident classroom environment are grouped so that they have approximately the same level of language proficiency, learning materials used at a typical nonresident facility like a Consolidated Training Center (CTF) computer center must be capable of addressing the needs of individual students who are at different levels of language proficiency. In a twelve-position facility, it is not uncommon to find ten different programs in terms of the level of difficulty of the materials being used.

(5) *Nonresident training programs allow for immediate feedback/review/advice in order to avoid habituated errors made by the learners.* This is perhaps more true in relation to foreign language instruction than other types of instruction since skill acquisition is so highly cumulative.

(6) *Extensive diagnostic treatments and placement procedures are absolutely essential in nonresident training programs.* Since formal contact with a teacher/tutor is sporadic in nonresident programs, every effort must be made to determine the ability of the learner to grasp instruction at a particular level of difficulty in any given program; in addition, every effort must be made to properly place the learner so that assigned tasks are attainable within a normal range of effort. More problems occur with student motivation because of

improper diagnostic and placement procedures than with any other single variable in the teaching-learning environment.

(7) *Methodology in nonresident training programs usually encompasses the 'mastery model'.* In a conventional learning environment, students progress as a group and are compared to each other as they develop language skills. In a nonresident learning environment, instruction is individualized to such an extent that students are not compared with each other; rather, their efforts are compared to previous performance and every attempt is made to master the materials at an acceptable level of proficiency.

(8) *Learning in a nonresident environment must be self-pacing for the student.* Nonresident courseware must be highly versatile since different students require alternate paths and varying amounts of reinforcement and review. Courseware versatility is the very foundation of nonresident learning programs.

(9) *The learner controls the instruction in nonresident training programs.* One of the major differences between a nonresident and resident training program is the degree to which students control the learning environment. Courseware must be packaged so that there are more options available to the learner. Students have different learning styles to the same extent that teachers have individual teaching styles. Courseware must be selected/developed accordingly.

(10) *Nonresident training programs must anticipate and deal with the 'forget factor'.* In a conventional classroom, teachers are effective to the extent to which they can anticipate/predict questions and problems encountered in that environment. In a nonresident training program, the courseware must be selected/developed based on the extent to which student questions and problems are covered effectively. The nonresident student simply cannot call a teacher/tutor at the CTF for every question/problem that may occur. Courseware must be almost totally self-contained; that is, most of the student's questions must be anticipated and addressed in the original courseware package.

(11) *Record keeping is an essential component of nonresident training programs.* Conventional classroom programs do not require extensive record-keeping procedures since the regular classroom teacher works with students on a daily basis and is usually well aware of deficiencies and dysfunctions. In the nonresident learning environment, there are so many students at varying levels of proficiency with unique configurations of capabilities and special needs that a sophisticated system of record keeping is absolutely mandatory.

Note. It is obvious upon reviewing the characteristics of nonresident language learning programs listed above that high technology is the best single solution to meeting the needs of this type of instruction. A mistake often made by developers of nonresident instructional programs is that they attempt to use course materials from resident training curricula. This approach does not result in satisfactory experiences because the nature of the nonresident learning environment is so different that needs of learners in this situation must be addressed with courseware that is packaged primarily with their specific job tasks and needs foremost in mind.

Computer-Adaptive Testing/Computerized Adaptive Language Testing (CAT/CALT). Computer-adaptive testing is a variation of tailored testing which permits the determination of the sequence of items encountered to be based on the ongoing pattern of success and failure experienced by the examinee. Most commonly, such an approach would, for an examinee who experienced success with a given item, result in the purposeful presentation of an item of greater difficulty. The examinee who experienced failure with a given item would next encounter an item of lower difficulty. Some variation of this process would continue in an iterative fashion until it was determined that sufficient information had been gathered about the ability of the examinee to permit the termination of the test.

Advantages of Computer-Adaptive Testing. Profound advantages accrue through the use of such an approach.

(1) *Individual testing time is reduced.* Since the test program is developed so that a given examinee is most likely to encounter only items matched to his or her ability, fewer items need be encountered than in the standard paper-and-pencil testing situation where every item in the test is presented, regardless of whether or not these items are matched to the ability of any given student.

(2) *Frustration and fatigue are minimized* in the case of students who would otherwise be forced to respond to quantities of items that would be below or beyond their ability levels.

(3) *Boredom is reduced* for those examinees who would otherwise be asked to encounter numbers of items that would be too easy for them.

(4) *Test scores are provided immediately* since scoring is performed by the computer as the exam is in progress.

(5) *Diagnostic feedback is given immediately* to examinees, teachers, and/or administrators.

(6) *Test security is enhanced* since it is highly unlikely that any two examinees would encounter the exact same set of items in exactly the same sequence.

(7) *Record-keeping functions are improved* so that it is more feasible to maintain longitudinal records of the progress of individual students than is the case with other testing programs.

(8) *Reporting, research, and evaluation capabilities are expanded* since detailed information is stored concerning the performance of each student and may be considered in relation to other demographic information about the student.

Satellite communications. The combination of satellite reception and videodisc applications allows students to role-play in a safe, simulated, authentic environment that places language in the proper context as the life-blood of the host culture and an outstanding tool for teaching culture. Contextual methodologists tell us that there is no substitute for access to foreign television, radio, and printed media--all produced by native speakers for comprehension by native speakers. This is true language in context.

CD-ROM/CD-I features and learner options. Some of the most significant developments in the area of Optical Disc technology are found in CD-ROM/CD-I applications. CD-ROM features include quality audio, static graphics, 550-600 MBs of storage, long life, and rapid random access. CD-ROM learner options include text, regular audio, slowed audio/more carefully enunciated audio, multilingual dictionaries and thesauri, maps and graphics, interactive reading and listening comprehension, audio with graphics, audio with text, word-look-up dictionary/thesaurus, and cultural notes. CD-I features include all features previously mentioned for CD-ROM plus full-motion video and split-screen displays. CD-I learner options include all those previously listed for CD-ROM plus AV with text, AV with static/animated graphics, AV with text and graphics together or separately, video and text displayed on partial/split screens, and updating maps paralleling video.

Judging from the tremendous growth experienced in CD-ROM technology in the United States alone, these latest applications of optical disc technology are here to stay. In the five-year period from 1987-1992, actual market figures and conservatively projected expectations in the areas of CD-ROM revenues, sales of CD-ROM drives, and sales of CD-ROM discs indicate that this represents by far the most rapidly growing application of technology to date.

Conclusion. Applications of technology toward the enhancement of the teaching and learning of languages have led to trends that have been very positive in their impact in this critical area of education. We have seen the importance of the universality of high tech applications, most notably found in the designation of interactive learning centers, computer centers, and testing centers rather than the traditional concept of the 'language laboratory'. This has enabled us to work cooperatively with other disciplines and with colleagues who have been able to help us reach goals including the exportability of our learning facilities as well as learning packets, the development of computer adaptive tests, the reception of satellite programs as well as satellite teleconference broadcasts, the more sophisticated development of paper as well as programming templates, and the growth of the concept of nonresident training used to support, maintain and enhance language proficiency. Many of these trends are possible because of the importance placed on the concept of a consolidated teaching and learning facility.

We have seen the expansion of goals in the areas of interactive simulation, program accountability and flexibility, increased student motivation as well as satisfaction, expanded modularity and time on task, and the genuine feeling on the part of faculty and students that our menu-driven, transparent CAI/IL programs are high tech, high touch, and friendly. We truly have made great progress and will continue to go forward at an increased pace with the help of CALICO and the many professional language and technology organizations that foster true cooperative effort.

Language learning, cultural understanding, and the computer

Judith G. Frommer

Harvard University

Introduction. Is cultural understanding necessary for language learning and is a knowledge of the language necessary for cultural understanding? A recent travel ad in the *New York Times* proposed '3 cultures for as little as \$899 in 8 days, including an English-speaking guide.' This ad implies a definition of culture that excludes language, that is based on architecture, geography, cuisine and art. It is a nuclear, hermetic, static, culture remaining on the surface of the physical and material and ignoring the human aspect, which is reached only through verbal communication. It reminds us of Jacques Tati's nearly silent film, *Playtime*, in which American tourists spend their week's vacation in the ultramodern quarter of *La Défense* on the outskirts of Paris, seeing the Eiffel Tower only in the reflection of the revolving doors or the exterior panels of enormous glass skyscrapers.

Both the advertisement and *Playtime* implicitly deny or ignore the cultural component of language, and the role that language plays in projecting cultural significance and fostering cultural understanding. None of us here would disagree, and yet we are often guilty of the opposite of the ad and *Playtime*--rather than trying to communicate cultural concepts without language, we teach language devoid of cultural concepts. We do this when we teach vocabulary independent of context, when we disregard semiotics, and when we neglect cultural referents essential for understanding the target culture.

We should perhaps pause here to define our terms: are we talking about culture with a capital *C* or a lower-case *c*, and what do we include in each category? Wilga Rivers, in *Teaching Foreign Language Skills*, quotes a report, published in 1904, in which the literature and the spirit of the nation were mentioned with regard to cultural understanding, whereas, she continues, the 1979 President's Report referred to unlocking the mysteries of other customs and cultures. Today, when language teachers speak of culture, according to Rivers, they are usually referring 'to all aspects of shared life in a community' (Rivers 1981:316).

Cultural links to language. Accepting this broad definition, I would like to propose that there are at least five ways in which culture is imbedded in language: linguistically, discursively, semiotically, intellectually and ethnically. First, linguistically, there are certain lexico-grammatical patterns and structures that are situationally appropriate because of custom. Native speakers of the target language react in a specific way to a specific rejoinder, and anything else would be inappropriate. A nontypical response would either be

considered impolite or create bewilderment on the part of native speaker interlocutors, because it would confuse them as to what their next move should be.

Discursively, culture determines discourse strategies and speech styles. This goes beyond the appropriate, automatic response discussed above. The degree to which ideas are presented directly or obliquely, the ways in which one takes turns in conversation, and the ways in which one makes requests or presents objections depend not so much on specific vocabulary as on attitudes and approaches (see Kramsch 1989:27-28 of original paper). One indication of these differences was the reaction of American students to a taped interview between an American and a French woman; they found the French woman impolite because she kept interrupting the interviewer, not realizing that in a discussion in France it is accepted behavior to anticipate the end of the other person's sentence and to begin speaking just before it ends without waiting for a pause, as is the custom in the United States.

Culture is also linked to language semiotically when words suggest meanings that go beyond the simple dictionary definition of the word. In this category are words and expressions that, rather than having a specific temporal or structural designation, evoke abstract ideas or concepts or characterize a society. For example, *le baccalauréat* is defined as a certificate attesting to the successful completion of secondary school, but to the French it represents a label of respectability without which, in the past, one could not hold up one's head in polite society and without which, today, one is doomed to a life of economic subsistence. On a less emotional level, the metaphors of a language are signposts of the interests of a society. French expressions--*la crème de la crème, le panier à salade, boire jusqu'à la lie*--include references to food, whereas American expressions--*an ace, three strikes and you're out, that's par for the course, an idea from left field, home safe*--refer more frequently to sports.

Intellectually, quotations, events, names, dates, and titles may refer to a heritage so common to all inhabitants of the target culture that they are used without explanation or convey ideas and meanings independent of their origin. This category includes all that is related to 'high' or 'low' culture, not only literature, cinema, history, and art, but also television, comic books, and even trademarks or children's games. In a normal conversation in French, one must expect references to Monsieur Jourdain or Combrai, while a college-educated American will drop names such as Gatsby or Uriah Heep. Actually, today's television generation (even from Harvard) might have more of a tendency to mention Mr. Rogers, Pebbles, or Cheers. An excellent illustration of the way a cultural background in the language is essential to true comprehension and communication can be found in a recent issue of *L'Express* in an article entitled *PPDA: les rêveries d'un présentateur solitaire* (see de Saint Pern 1989). This title combines references to high and low culture. First, an assumption that the reader is familiar with French television is made by the author, since one can read the entire article about *PPDA*, the French equivalent of Dan Rather or Peter Jennings, without learning his name: Patrick Poivre d'Arvor. Also, without a knowledge of French literature, one would not recognize that the subtitle of the article, 'les rêveries d'un présentateur solitaire', is a pun on *Les Rêveries d'un promeneur solitaire*,

a book of preromantic essays by Jean-Jacques Rousseau. This type of information is necessary to carry on a meaningful conversation in the target language.

Finally, ethnicity also affects communication. Conversations do not take place in a vacuum, but rather between individuals from diverse environments that influence their perception of the words of their interlocutors. Unfortunately, stereotypes and xenophobia exist in all countries, so that the interpretation of a message often depends on a preformed impression of the group to which the interlocutor belongs and, therefore, a preconceived notion of his or her mindset. This is especially important for American students to understand since, often without realizing it, they are seen as the representatives, and sometimes even the symbol, of a capitalistic, imperialistic government. Words or gestures that seem innocent when expressed by a non-American, may provoke a hostile reaction in the TL country when spoken by a citizen of the United States.

The fact that language reflects culture, as indicated above, implies that one cannot truly learn the language without also learning about the culture; communicative competency is more than ordering food in a restaurant or asking directions to the train station. According to Jay, quoted in Seelye (1984:21): 'With knowledge of the language must exist a similar knowledge of the social, religious, and economic attitudes of a people ... Unless the student is learning the language in the target culture, the cultural referents necessary to understanding a native speaker must be learned in addition.' Seelye (1984:25) concludes: 'In short, it is becoming increasingly apparent that the study of language cannot be divorced from the study of culture, and vice versa.'

The advantages of the computer in general as a culture-language teaching tool. Accepting the inclusion of culture as a component of language courses, we are faced with the challenge of presenting cultural material to students so that they can best assimilate and appreciate it. The traditional methods for including culture in language-learning, such as exposition and explanation, especially of factual information, will often bore students. Skits and role-playing, while enjoyable and effective methods and essential for speech production, are limited and can become repetitive. In contrast, the computer provides a process for bringing culture to students in a variety of ways, allowing certain aspects of a culture to be highlighted while students execute certain operations.

The way the computer handles the teaching of situationally appropriate responses will differ from the ways in which it will be used to introduce the student to the cultural environment of the target language, but regardless of these variations, the computer has characteristics that make it particularly suitable as a facilitator of culturally aware language acquisition. First is the inherently interactive nature of the computer; passive reception of computer delivered material is impossible since the computer progresses only in response to user action. This characteristic is especially important with regard to the use of authentic documents, either in written, aural (tape) or visual form (graphics and video images) now accepted as essential materials for language learning. In practice, reading, listening, and viewing can easily

become passive activities and authentic documents can often be difficult or hermetic. Even though we, as teachers, usually find such authentic documents fascinating, it is often because of our ability to comprehend the language and the semiotics; knowing the signs and symbols we are looking for, we read texts, listen to tapes, or watch a videotape in an analytic way to assure capturing the semiotic message and feel a sense of pleasure when we do. In order for our students to react as we do, the activities must be made active and comprehensible. Most teachers tend to do this with study sheets or advice to use a dictionary, often with less than satisfactory results. Computer-presented texts, or interactive audio or videodisc, controlled by the computer, make it impossible for students to doze while supposedly reading, listening, or watching. The computer assures that they are active and, with appropriate software, can attract their attention to those culturally relevant details that we feel are important.

The second advantage of the computer is that it permits a multiplicity of experience impossible in the context of the normal academic program at any level. Students can work one-on-one with a computer to see more material or to have more practice for an unlimited period of time; although computers do occasionally crash, they never get tired, and they have infinite patience. Individualized instruction, impractical and unavailable in most schools and universities, can become a reality with computers. More importantly, the computer can provide a private, tension-free environment controlled completely by the student.¹ In addition, when students work together, the computer becomes a focal point for student interaction, independent of the teacher. This is another way in which use of the computer is an active rather than a passive experience; even in the most student-centered classrooms, the teacher never completely relinquishes control, and students are always aware of his or her presence.

In general, we can say that the computer contributes to the effective integration of culture into teaching and learning materials because of its ability to control, limit, specify and highlight. When teaching culture in language courses, there is the danger of drowning students in too much cultural information, which they are not prepared to absorb, or giving the bare minimum that leads to the unfortunate formation of stereotypes. The computer helps to avoid these pitfalls by depending on the wishes of the teacher, controlling the amount of material that is presented to the students, allowing them to explore vast cultural data bases, drawing their attention to details of culture that are small but significant, and by coaching them to react with the culturally appropriate response. At the same time, the immediate feedback aspect of the computer prevents misunderstandings from occurring initially. Most importantly, the constant interaction demanded by the computer maintains student interest. The computer has the added advantage of neutrality, and can introduce elements of culture independently of value judgments; we teachers cannot always claim to do so.

Examples of specific ways in which the computer aids in cultural understanding. Even without an interactive videodisc interface, which will be discussed below and which is more powerful, the computer alone, using simple software that does nothing more than present text on the screen, can

be used to promote culturally based language learning. Simulations, presented in the form of dialogs or descriptions of situations and followed by multiple-choice questions, can be effective. To help students to internalize the automatic responses necessary in any culture, the computer can present typical conversations that end with a rejoinder by one person and ask the student to choose the appropriate response from a list of three or four. If students are given meaningful error messages and allowed to use the same expressions in a variety of contexts, they will assimilate the culturally appropriate response. Although there is a tendency in the current atmosphere of proficiency-oriented language courses to disregard the importance of memorization and repetition in language learning, empirical evidence suggests that habit formation is an important aspect of language learning and automatisms are recognized by many as important in cognitive processes.²

Students can also use text-based software to learn how a target language speaker would describe a certain situation. The computer could present three or four descriptions and ask the students to choose the one that most reflects the speech styles of the target language. After choosing they would receive explanations for why the incorrect choices are culturally inappropriate and why the correct choice reflects the culture of the TL country. Another possibility would be to present students with a speech passage and ask them to choose among a number of explanations in order to see if they have interpreted the message according to the signals imbedded in the target language.

The examples given thus far are in the category of product-oriented learning in which the computer directs the students to perform certain tasks and verifies their performance. Completely text-based computer applications also serve to encourage process-based learning. Students can use software that allows them to explore data bases to investigate whatever they wish to, or whatever their teachers want them to, in order to learn more about the culture of the TL country. This is applicable to the semiotic and intellectual aspects of cultural understanding, the necessity of knowing what the important names and dates stand for and how target language speakers react to them, and the importance of possessing a knowledge of the cultural heritage of the target language country. If one thinks of foreign students learning ESL, *Cultural Literacy* by E. D. Hirsch (Hirsch, Kett, and Trefil 1988) is not really useful, since, in the first edition, all one sees is a list of words with no explanation as to their meaning, usage, or relationship to each other. For example, 1984 is listed as an important date, with *title* in parentheses; George Orwell's name is included in the list, but there is no indication that the two are related. Imagine this same list on the computer: by selecting any entry, either by clicking on a mouse or touching the screen, one could see an explanation including both text and graphics and could explore related items in the list. Selecting the name of an author included in the list could produce a succinct biography, a summation of his work, and a list of titles. The user could then select one of the titles and find out bibliographical information, learn which libraries have which books, and even read passages or an entire work.

Critics might say that, using the computer as described above, students will become lazy and incompetent, that they will no longer learn how to use

a library or to do research, and that they will never browse and make a discovery of unexpected but valuable information. I think just the opposite is true, and for at least three reasons. First, if we consider the example given above, identification of the items in the *Cultural Literacy* list, those who do not recognize the items on the list would probably not have the least idea of how to go about finding out about them. Second, students rarely make or have the time that it takes to physically find the information that can be made available instantaneously in computer-provided relational data bases. Finally, anyone who has used a computer data base knows that browsing is possible on the computer; it is so easy to click on the mouse or to touch the screen or type a few symbols that one is very easily drawn into extensive exploration on the computer in the same way as one picks up book after book in the library stacks. With extensive data banks available, 'A twenty year old sophomore can now put his or her hands on more information than the most learned scholars a few years ago could ever collect.'³

Especially at beginning and intermediate levels, the electronic computer-assisted data base exposes language learners to cultural information to which they would not otherwise have easy access. In fact, the criticism of *Cultural Literacy: What Every American Should Know* could be countered by a computer application. If the meaningless list were transformed into a data bank which allowed the user to establish his or her own links as in Intermedia, it would become a true learning experience.⁴ The sorting capability of the computer could also be used to help students to understand culture. Consider a linguistic data base or on-line dictionary from which students could call up notional-functional lexicons. Typing 'introductions' would produce a list of all of the different ways in which people can be introduced in the target language. Above, we cited the word *baccalaureat* to illustrate how culture is linked to language semiotically. A sorting program could find all words associated with education that are especially important to the French, and could present them to the student-user with an explanation and examples of usage.

The visual component. The possibilities of the computer to promote cultural understanding are magnified infinitely by the addition of images, whether static or animated graphics, or interactive videodisc. The importance of providing a visual environment for language learning cannot be overestimated. Halliday (1978:61) emphasizes the background information necessary for an individual to understand a message:

From a sociolinguistic standpoint, a text is meaningful not so much because the hearer does *not* know what the speaker is going to say ... but because he *does* know. He has abundant evidence, both from his knowledge of the general ... properties of the linguistic system and from his sensibility to the particular cultural, situational and verbal context; and this enables him to make informed guesses about the meanings that are coming his way.

In fact, students who have never seen the target language country nor interacted with native speakers do not have the knowledge of the 'particular

cultural and situational context' mentioned by Halliday, and that is why they have difficulty making 'informed guesses.' I would suggest that the absence of this context is a factor affecting the differences in native, second and foreign language acquisition. Normally, we expect our students to become proficient in a language speaking it in an alien environment; by adding an authentic visual component, we can provide them with the nonverbal information they need to mentally construct the appropriate environment for assimilating the language.

The visual component provided by the computer can be in the form of static graphics, animated graphics, or video images. Even though the computer graphics are not as vivid as the video images, they still offer a visual stimulus that aids understanding and retention, on the condition that they are high-quality and integrated into the software. They can be extremely useful as illustrations of objects or gestures in first-year courses, especially for non-Roman alphabet languages. Graphics can establish a realistic visual context for a game, such as solving a mystery, or an obviously cultural activity, such as finding one's way around a city in the TL country. This type of software is within everyone's reach with microcomputers and inexpensive scanners.

Interactive videodisc. While graphics are effective, they are not so realistic or absorbing as the video images produced by videodiscs. The marriage of the visual impact of the videodisc with the technical power of the computer produces an exposure to culture that the student could not otherwise experience outside of the TL country. It has been said that 'an interactive videodisc system is not just a merging of video and computer media, but that an entirely new medium arises with characteristics that are different from each of the composites' (Verano 1989:249).

At this point, it might be useful to mention that a videodisc consists of 55,000 individually numbered frames and that showing a sequence of frames is like showing a film sequence, while showing a single frame is like showing a slide. Programs that control videodiscs use the frame numbers to present specific segments or images to the viewer who has only to react, usually by clicking on a mouse, typing a simple command, or touching the screen. Depending on the hardware, the videodisc image may be seen on the computer terminal or on a separate monitor.

Students can profit more from interactive videodisc than from simple videotape, because the latter provides only a passive activity; as one college English professor stated: 'I don't believe people really *think* when they watch video.'⁵ Using interactive video is by definition, as already mentioned, an active experience. The most clear-cut advantage of interactive videodisc over videotape is that interactive videodisc offers nonlinear viewing. Because a computer-controlled videodisc can be accessed randomly, students can be shown or choose to see segments in any order and can, as a consequence, create their own stories from a group of scenes, refer back to scenes already viewed, and be directed to special segments for remediation. In addition, segments can be replayed to highlight their cultural relevance.

Specific examples of interactive videodisc projects. Let us now consider specific examples of how interactive videodiscs help students to understand the culture of the target language country and to acquire cultural as well as linguistic competency in the language. Two existing videodisc projects, although not created for language learning, demonstrate the possibilities of this medium. In the first of these, developed at Harvard Law School, students view a trial simulation and at any time can interrupt the videodisc to ask questions or to raise an objection to the procedures. When students object, the program tells them whether or not it is legally possible to make an objection at that time.⁶ The second videodisc project, based on the technique used in the Law School videodisc, is used for personnel training; it presents simulations of interviews in which we see and hear an interviewer talking with a job applicant.⁷ At any point, the user can interrupt the videodisc and question the appropriateness of the interviewer's or interviewee's remarks or comportment.

This type of approach could easily be used for foreign languages. Students could watch a conversation with the capability of interrupting at any point to ask why certain things have been said or, in the case of a conversation between a native speaker and a foreigner, to indicate what the foreigner has said incorrectly and why. This could even be set up as a game in which students earn a point for every correction made and lose a point every time they try to correct a perfectly acceptable rejoinder. A variation of this would be to have the student choose among a few responses at a particular point. A student who has chosen an inappropriate response can be shown a segment that demonstrates the consequences. For example, if a student of French, viewing a scene at a dinner table, chose to say *Oui, merci* 'Yes, thank you', the French translation of an appropriate American reply, when asked if he or she wanted a second helping, the following scene could show him or her being passed over, thereby visually emphasizing the point that saying 'thank you' means that you are refusing, and that people focus on that word and often do not even hear the 'yes' that precedes it.

Another application of interactive videodisc is to offer students the possibility of focusing on the visual image by having them look at a single frame. For example, in the BBC French videotape series, *La Marée et ses secrets*, the first scene shows a young man reading a newspaper, *Ouest France*. If this videotape were on videodisc, the student could see a still frame of the newspaper and notice the title. Sophisticated software could allow the student to click on the newspaper to learn about its tendencies, origins and circulation, and maybe even to read it.

Single frames can also be based on slides that have been transferred onto a videodisc. The videodisc of the Perseus project, a complete multi-media course on Ancient Greece, contains 2,100 images based on slides, 1,800 of sites and monuments and 300 of objects.⁸ Using the computer, the students can access any of these slides and, at the same time, see an explanation on the screen. The Perseus videodisc also contains twenty minutes of motion video consisting of narrated descriptions of the major sites. In addition, the program includes on-screen graphics that are interconnected so that students can select a city on a map, find out what happened there, read texts about it, and see artifacts from the period. Although based on ancient Greece, it is

evident that the Perseus approach could be applied to creating computer-delivered cultural components of modern language courses, especially in the area of intellectually based cultural understanding.

The MIT Athena Project French interactive videodisc project, *Direction Paris*, consists of two videodiscs, one a fictional account of the life, labor, and love of a young Frenchman, and the other a documentary based on a neighborhood in Paris.⁹ Since both of these were filmed in Paris, they obviously provide authentic cultural materials. The originality of the fictional account, entitled *A la rencontre de Philippe*, is that it is nonlinear; after an initial scene, student decisions at crucial points determine how the story evolves. The story can be affected by students' answers to specific questions, based on comprehension of a specific segment, or by a decision that results in omission of a segment.

There are at least two ways in which *Direction Paris* illustrates how computer-controlled interactive videodiscs can serve to promote cultural understanding. First, the obstacle to students' appreciation of most authentic materials, difficulty of comprehension, is overcome by the capacity of the computer to provide on-line explanations and translations. Students have two alternatives when they do not understand: (1) they can listen to an alternate sound track with completely natural but more clearly enunciated French; or (2) they can stop the videodisc at any point and obtain an explanation of what they have just heard in the form of a summary of the action or the corresponding passage of the French script. Although a second sound track can also be made available on a simple videocassette, using it is much more difficult for the student because the inaccuracy of the counter on a VCR makes it difficult to review specific tape segments. Providing the script is also quite different from showing videotapes with subtitles that are always present on the screen blocking the students' perception of the visual content and serving as a crutch (see Massadier-Kenny 1988). In order to access the script of a computer-controlled videodisc, students must consciously perform an action (clicking a mouse, touching the screen or hitting a key) to stop the video and display the text. Thus, when they are viewing the videodisc they are not distracted by subtitles and can concentrate on the total message transmitted by the words and the image as well.

Second, the nonlinear approach of *Direction Paris*, possible only with videodisc technology, immerses the students in the culture of the language they are studying. In the documentary videodisc, the student can select a street in the neighborhood in which the story takes place, then select a shop, read about that shop, and finally see a video interview with the shopkeeper and customers. The technological capacities of the videodisc allow the student to link the various cultural components, providing a coherent context and, therefore, greater understanding and retention. In *A la rencontre de Philippe*, the student is invited from the very beginning to participate in the action. First, there is a specific invitation to come to the café. Then, Philippe asks the student to help him, and assigns the specific task of going to his apartment to listen to the messages on his answering machine. The student is able to go to the apartment and to walk through it room by room. He or she can then find the answering machine, push the buttons, listen to the messages and write them down. Then when Philippe asks for the number the student can

type it and hear the phone being dialed. This immersion in the action simulates real life in France and the student absorbs small details of language and culture naturally, rather than having them delivered didactically.

My examples thus far all refer to student use outside of class. However, the possibility of accessing an exact location on a videodisc, either a segment or a single frame, with a computer can aid teachers in presenting culture to students. Any teacher who has tried to show an exact spot of a videotape to a class in order to illustrate a point, knows how much time is wasted, with inevitable student distraction. Simple software exists that allows teachers to prepare effective classroom demonstrations of culturally relevant materials that will hold students' attention.

Interactive audio. Although less impressive and accurate than interactive videodisc, computer-controlled tape recorders should not be overlooked as an aid to improving students' cultural understanding. Possessing many of the same characteristics of interactive videodisc, the computer-controlled tape recorder can replay tape segments or provide transcripts of difficult material, either automatically or based on student request. Many authentic sound documents, often too difficult even for advanced students, can become accessible with interactive audio (see Jones 1989:219-33).¹⁰

Electronic communication. The computer can be used, in an entirely different way from those already mentioned, to provide students with a completely authentic cultural experience. Although it is not always easy to arrange and necessitates extensive technical support, the existing electronic computer network can be used to put students in direct contact with individuals in the TL country. What better way to learn about culture than through direct communication?

Conclusion. The computer, then, possesses definite advantages in introducing students to meaningful cultural understanding through language learning. Student control, immediate access to data and images, the possibility of using a nonlinear approach, and the capacity to interface more than one medium all contribute to the computer's effectiveness as a facilitator of language acquisition. Also, the computer, used outside the classroom, can provide additional exposure to culture and can offer extra help outside of class hours; the computer allows culture to be assimilated not only better, but faster. Amazingly, little of the software described here existed even five years ago, and five years from now much more will be possible because of advances in computer technology.

The computer will not realize its full potential, however, until certain problems are solved. Of a practical nature are the aspects of funding and training of teachers. While these problems are considered primordial, in fact the expenses of purchasing hardware and software and teaching teachers to use them, although high, will not prevent the extension of computer and interactive videodisc usage. A more serious problem is the availability of quality software: much of what I have described is what could be done, not what actually exists. The superb Perseus and Athena project packages will soon be available, but they represent two languages, each at a specific level.

To create materials at all levels in all languages, it is not only necessary to obtain funding but, more importantly, to find original thinkers who will spend the time to develop them. Given the realities of course loads and tenure track obligations, such materials will result only from a national effort, not the usual approach in the United States. Finally, the most serious obstacle to the successful integration of technology into language courses as a conveyor of culture is a lack of knowledge about implementation of courseware. It is evident that the computer can help make our students culturally aware, but we must provide the materials, learn how to use them effectively, and stimulate our students to want to use them.

Notes

1. The environment will only be tension-free if students feel comfortable using the computer. All computer use should be preceded by orientation sessions to ensure student understanding, not only of computer operation, but also of the learning goals of the computer activity.
2. For example, Victor Shklovsky wrote: 'If we start to examine the general laws of perception, we see that as perception becomes habitual, it becomes automatic.'
3. Gregory Crane et al. in the Perseus Project proposal, January 1986, p. 13.
4. Intermedia is a software project being developed at Brown University. Intermedia creates an environment in which data can be linked in many ways by the creators or users of data banks.
5. Richard Marius, quoted in an interview in the *Harvard Gazette*, March 3, 1989, p. 5.
6. Harvard Law School Project, an ongoing videodisc project currently used by students.
7. The personnel videodisc is the work of David Stein, Office of Patents and Copyrights, Harvard University.
8. The Perseus Project, a joint venture of Harvard University and Pomona College, among others, was funded by the Annenberg CPB project. The first version will be released for distribution in late 1989.
9. Gilberte Furstenberg, lecturer at MIT, is the author and developer of *Direction Paris*, in the context of the Athena Project, under the direction of Dr. Janet Murray.
10. A number of authoring systems exist for preparing interactive audio courseware among which is the *MacLang Authoring System* for use with the Macintosh and the Tandberg tape recorder.

References

- de Saint Pern, Dominique. 1989. Les rêveries d'un présentateur solitaire. *L'Express*, 13 January 1989, 36-37.
- Halliday, M.S.K. 1978. *Language as social semiotic*. Baltimore: University Park Press.
- Hirsch, E.D., Jr., Joseph Kett, and James Trefil. 1988. *Cultural literacy: What every American needs to know*. New York: Vintage Books.

- Jones, Randall L. 1989. Using interactive audio with computer assisted language learning. In: W.F. Smith, ed., *Modern technology in foreign language education: Applications and projects*. Lincolnwood, Ill.: National Textbook Company. 219-33.
- Jones, Randall L., George M. Henry, et al. 1989. FLIS: Random-access and innovative lesson types. In: W.F. Smith, ed., *Modern technology in foreign language education: Applications and projects*. Lincolnwood, Ill.: National Textbook Company.
- Kramsch, Claire. 1989. New directions in the teaching of language and culture. In: *National Foreign Language Center Working Papers*, ed. R. Lambert. Washington, D.C.: NFLC. pp. 27-28 of original paper, presented at Bellagio, Italy, 1988.
- Massadier-Kenny, Françoise. 1988. La Vidéo: Mode d'emploi. AATF National Bulletin. 14.1.10-12.
- Rivers, Wilga M. 1981. *Teaching foreign language skills*. 2nd ed. Chicago: University of Chicago Press.
- Seelye, Ned. 1984. *Teaching culture: Strategies for intercultural communication*. Lincolnwood, Ill.: National Textbook Company.
- Sjlpvsdu, Victor. 1988. Art as technique. In: *Twentieth century literary theory*, by K. M. Newton. New York: St. Martin's Press. p. 23.
- Verano, Miguel. 1989. USAFA interactive study in Spanish. In: *Modern technology in foreign language education: Applications and projects*. W. F. Smith, ed. Lincolnwood, Ill.: National Textbook Company.

New developments in knowledge-based machine translation

Sergei Nirenburg
Carnegie Mellon University

Recent experience in Knowledge-Based Machine Translation (KBMT) suggests that systems adhering to this paradigm can be useful not only in very small domains and limited sublanguages. With special attention paid to acquisition of large knowledge bases and with the advent of new tools (including representation languages, human-computer interfaces, database management systems, etc.) the practicality of the knowledge-based approach is steadily growing. Since a comprehensive automatic analysis of meaning is not yet feasible, and the attainment of this goal will remain the central objective of computational linguistics for years to come, a practical KBMT system will be necessarily of a hybrid nature. It may include certain types of processing that would be considered more appropriate for a transfer-oriented system. It will also include a measure of human involvement. However, while in transfer systems human involvement invariably means postediting, human-aided knowledge-based systems will use human help *during* the process of translation, to finalize the representations of the meanings in a source language (SLG) text. It is expected that the target language (TL) texts produced from such improved meaning representations will be of comparable quality with translations produced by humans.

1. 'Transfer or interlingua?'--Is this question still relevant? Historically, machine translation (MT) systems have been of three major types: direct, transfer, and interlingua. Detailed descriptions of the three approaches, with all their modifications and varieties, can be found in the MT literature (see, in particular, Hutchins 1986, Zarechnak 1979). Direct systems have been justly criticized for their *ad hoc*ness, so that at present the choice of architectures for machine translation systems is reduced to the two latter approaches. In this section we will very briefly comment on the essential differences between them and suggest that the latter are possibly less important methodologically than the attitude to meaning analysis and also coverage.

Transfer systems involve a measure of target language-independent analysis of the source language. This analysis is usually syntactic. It allows substituting SLG lexical units with TL lexical units *in context*. That is, it permits taking into account the types of syntactic sentence constituents in which lexical units appear.

In interlingua systems the SLG and the TL are never in direct contact. The processing in such systems has traditionally been understood to involve

two major stages: representing the meaning of a SLG text in an artificial formal language, *interlingua*, and then expressing this meaning using the lexical units and syntactic constructions of the target language. Few interlingua systems have been fully implemented because of the very significant complexity (both theoretical and empirical) of extracting a 'deep' meaning from a natural language text.

The major distinction between the interlingua- and the transfer-based systems is, in fact, not so much the presence or absence of a bilingual lexicon but rather the attitude towards comprehensive analysis of meaning. In practice, those MT researchers who believe in translating without 'deep' understanding (or perhaps who believe in the unattainability of 'deep' understanding) of the SLG text tend to prefer the transfer paradigm. The price they have to pay for avoiding meaning analysis is the need for an extra step in the translation process, namely, postediting.

Inherently, a transfer system can involve many levels of meaning analysis. This becomes clear when one considers that different transfer-based systems have widely varying levels at which transfer occurs--from simple phrase structure trees to detailed representations that use subcategorization patterns, and even selectional restrictions. There is a trend in transfer-based MT to downplay the necessity of structural transfer, that is, the stage of transforming standard syntactic structures of the SLG into the corresponding TL structures. This is in part due to the prevalence of grammatical theories that eschew transformations and seek universally applicable representations of grammatical structures and relations. This trend is essentially interlingual in nature. Transfer-based systems can also deal with lexical semantics; the language in which the meanings of SLG lexical units are expressed will be the TL itself. This can be implemented through a bilingual lexicon featuring disambiguation information.

In interlingua systems the meanings are represented in an artificial language--the reason being that such a language is better suited for the formulation of disambiguation rules necessary for producing an adequate meaning of a SLG text, in part because it was specifically designed for this purpose.¹

Distinctions between the transfer and the interlingual approaches are best drawn at an abstract level. In reality, when practical systems have to be built, many types of work will be practically identical for both approaches (notably, the grammars and programs for syntactic analysis and synthesis). For some other types of work the very nature of the material dictates the necessity of methodological compromises--for instance, some source language lexical units for which the interlingua does not, at the moment, have an adequate representation can be treated in a transfer-like manner in a practical knowledge-based machine translation (KBMT) system. At the same time, for those (very frequent) cases when there is no possibility of direct transfer of a lexical unit or a syntactic structure between two languages, a transfer system would benefit by trying to express the meaning of such lexical units and syntactic structures and construct a TL correlate from this (more detailed and transparent) representation. The requirements of practical use, indeed, pose similar difficulties for both approaches--consider such universal problems as ill-formed input, special symbols and codes, document layout preservation,

translatable material in figures, etc. ATLAS-II (Uchida 1987) is an example of a hybrid MT system that has features from both major approaches.

2. New features in KBMT. I will illustrate the recent progress in knowledge-based machine translation using as the example KBMT-89 (Nirenburg and Goodman 1989), a system recently developed at the Center for Machine Translation of CMU. It translates from English into Japanese and from Japanese into English in the domain of computer hardware installation manuals. Small-scale extensions are being developed at present to add French, Russian, and Polish to the list of source languages. The system development was sponsored by IBM.

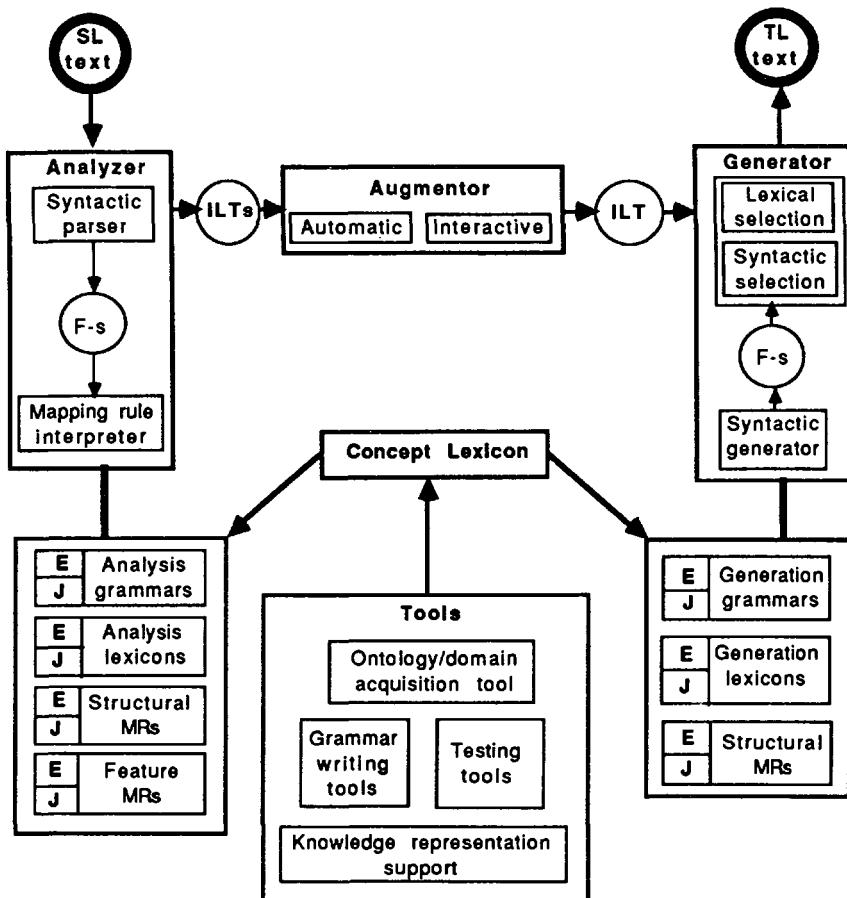
KBMT-89 consists of a large number of modules, as illustrated in Figure 1. There are four grammars; two analysis lexicons and two generation lexicons; a domain model (called ‘concept lexicon’); an augmentor, which serves as a filter between analysis and generation and a set of acquisition, maintenance, and testing tools. In terms of coverage and inventory of computational and knowledge components, KBMT-89 appears to be the largest KBMT system to date. For a sketch of the global architecture of the system, see Figure 1: The global architecture of KBMT-89.

In what follows I will briefly discuss several of the distinguishing features of KBMT-89. The system has many more salient features. See Nirenburg and Goodman 1989, for a more detailed discussion.

2.1 Nature and size of knowledge bases. KBMT-89 is supported by a domain model of about one thousand concepts organized into a complex network. As can be seen from Figure 2, the knowledge base used in KBMT-89 is capable of representing not only general domain knowledge about taxonomies of object types (e.g., ‘a car is a kind of a vehicle,’ ‘a doorhandle is a part of a door,’ ‘artifacts are characterized (among other properties) by the property *made-by*’) but also knowledge about particular instances of object types (e.g. IBM can be included into the domain model as a marked instance of the object type ‘corporation’) as well as instances of (potentially, complex) event types² (e.g. the election of George Bush as President of the United States is a marked instance of the complex action ‘to-elect’). The ontological part of the knowledge base takes the form of a multihierarchy of concepts connected through taxonomy-building links, such as *is-a*, *part-of* and some others. We call the resulting structure a multihierarchy because concepts are allowed to have multiple parents on a single link type.

In KBMT-89 the ontological concepts are first subdivided into objects, events, forces (introduced to account for intentionless agents) and properties. Properties are further subdivided into relations and attributes. Relations are mappings among concepts (e.g. *belongs-to* is a relation, since it maps an object into the set {*human *organization}) while attributes are mappings of concepts into specially defined value sets (e.g. *temperature* is an attribute that maps physical objects into values on the semiopen scale ‘> 0’, with the granularity of one degree Kelvin). Concepts are represented as frames whose slots are properties fully defined in the system.

Figure 1. The global architecture of KBMT-89.

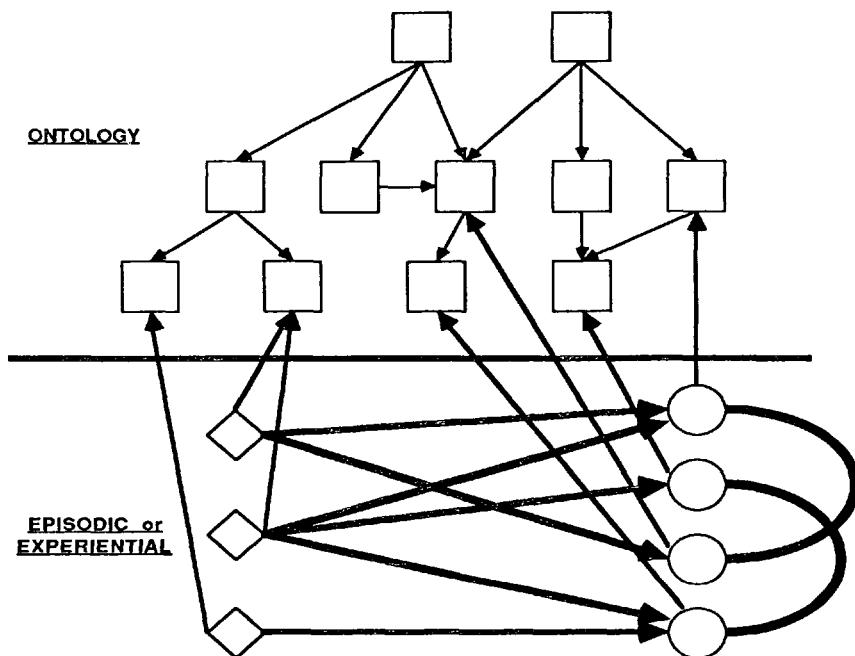


In order to acquire and maintain the KBMT-89 domain model we developed an interactive knowledge acquisition and maintenance system, ONTOS (for a detailed discussion see e.g. Nirenburg et al. 1988a). To illustrate the way ONTOS operates, consider Figures 3 and 4. They show the ONTOS screen at two strategic moments during the acquisition of the concept of the Macintosh microcomputer. The acquisition is performed by using and extending the information already available in the domain model about the parents and (if available) siblings of the new concept.

Many models of a single domain are possible, and it is methodologically improper to insist on there existing a single canonical domain model. The set of ontological postulates used in KBMT-89 (and illustrated here) has been deliberately made as general as possible in order to make it adaptable to other views of the world. Using the KBMT-89 acquisition tools and, optionally,

using the KBMT-89 domain model as the starting point, other researchers can build their own domain models within a short period of time--the task certainly impossible before the advent of knowledge acquisition and maintenance systems.

Figure 2. Metatypes of entities and relations in KBMT-89 domain model. The domain model also serves as an index into the analysis and generation lexicons for both English and Japanese. It represents both ontology (semantic memory) and experience (episodic memory).



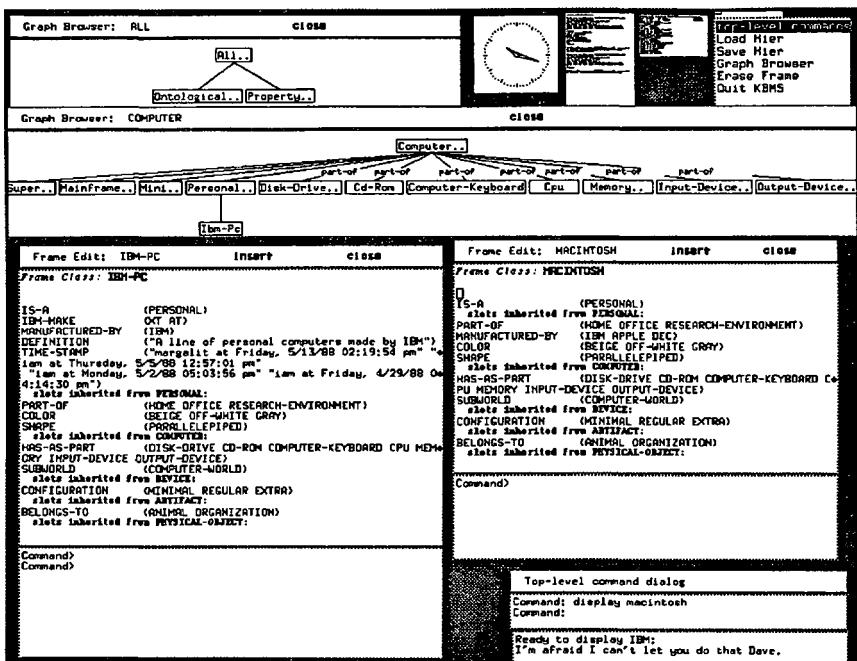
OBJECTS

- Ontological concepts
- Remembered Instances of objects
- Episodes (remembered experiences)

LINKS

- Ontological links: IS-A, PART-OF, etc.
- Instantiation (IS-TOKEN-OF) links; for indexing properties
- Semantic role links : AGENT, SOURCE, etc.
- Episode links: TEMPORAL, SPATIAL, CAUSAL, etc.

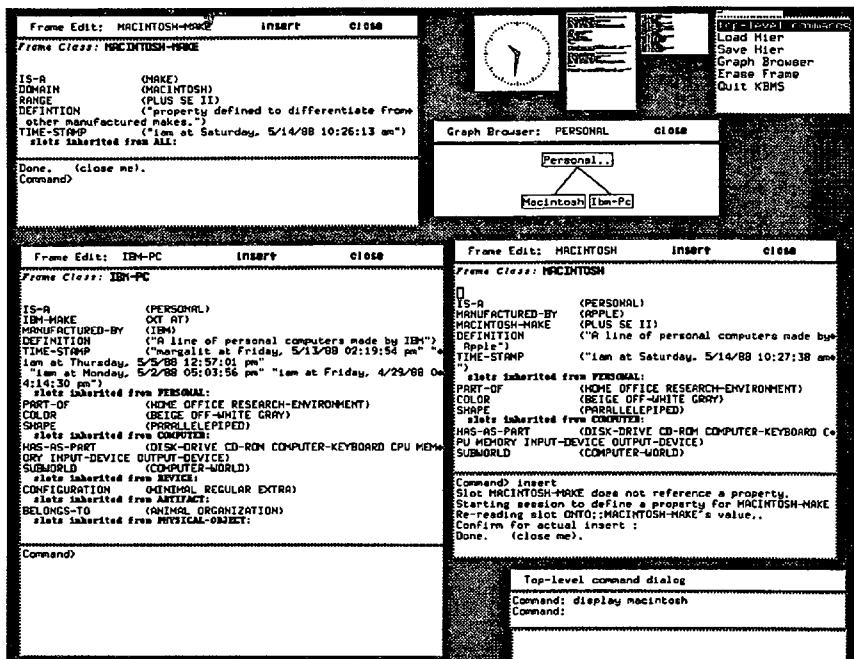
Figure 3. The concept 'macintosh computer' in its initial form.



2.2 Exploring additional types of meaning. Traditionally, the types of meanings represented in knowledge-based natural language processing systems have been almost entirely propositional. A large number of nonpropositional-pragmatic and discourse-meanings, such as thematic structure, speech act, modality, discourse cohesion, speaker attitude, etc. were not overtly represented. In knowledge-based machine translation such meanings have been traditionally ignored. In transfer systems they are treated implicitly, based on the observation that often no special processing has to be done other than simple transfer of lexical clues or, sometimes, a structural transformation. This approach is prone to error and does not support any paraphrasing capability necessary when text-level stylistic decisions are treated.

We have argued elsewhere (Nirenburg and Carbonell 1987) for the necessity of incorporating nonpropositional knowledge into the representation of the results of text analysis (known as interlingua text in KBMT-89). Figure 5 illustrates how both propositional (domain model-dependent) and nonpropositional knowledge are integrated in a single representation. Analysis lexicon entries are correspondingly classified into those mapping into instances of domain model concepts and those signifying nonpropositional properties and thus mapping directly into specific property values in ILT.

Figure 4. Adding properties on the basis of inheritance and sibling differentiation.



2.3 Focus on generation. Unlike most machine translation systems, KBMT-89 pays a significant amount of attention to the generation side of the process. To give just one example, let me illustrate one component of the generation process--lexical selection. When the knowledge-based approach is used, it becomes possible to enhance the process of lexical selection (lexical synonymy resolution) in generation. Figure 6 shows lexical selection steps used in KBMT-89.³ Note that the lexical selection process involves filters that are essentially syntactic and source-language dependent (such as subcategorization) as well as semantic filters (such as the meaning matching metric, which operates on language-independent meaning representations) and stylistic filters (for instance, the preference, while generating English, for a verb to realize the meaning of the head of a proposition).

2.4 Human-computer interaction. The idea of human-aided machine translation occurred to MT researchers very early. Of a number of ways in which humans can facilitate the process of automatic translation we are mostly interested in having a human user verify, improve, and finalize the system's decisions during analysis. The system may be unable (have no knowledge) to prefer one candidate reading of the input over another. Or, alternatively, its

knowledge may rule out all of the candidate readings. Human intervention may become necessary.

Figure 5. The interaction among lexicons and ILT. Note that some source language lexical units are connected to their interlingua meanings directly, bypassing the Concept Lexicon. The figure also illustrates the lack of symmetry in the treatment of lexical semantics in analysis and generation; the main problem in analysis is polysemy, while in generation it is synonymy.

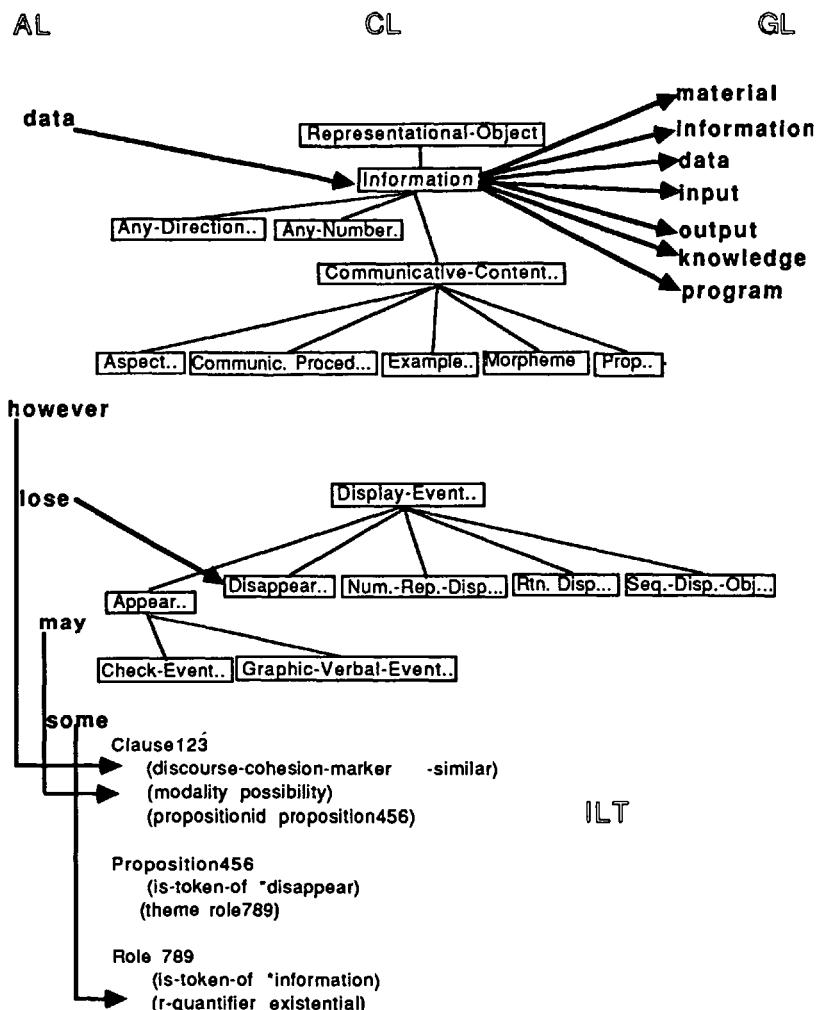
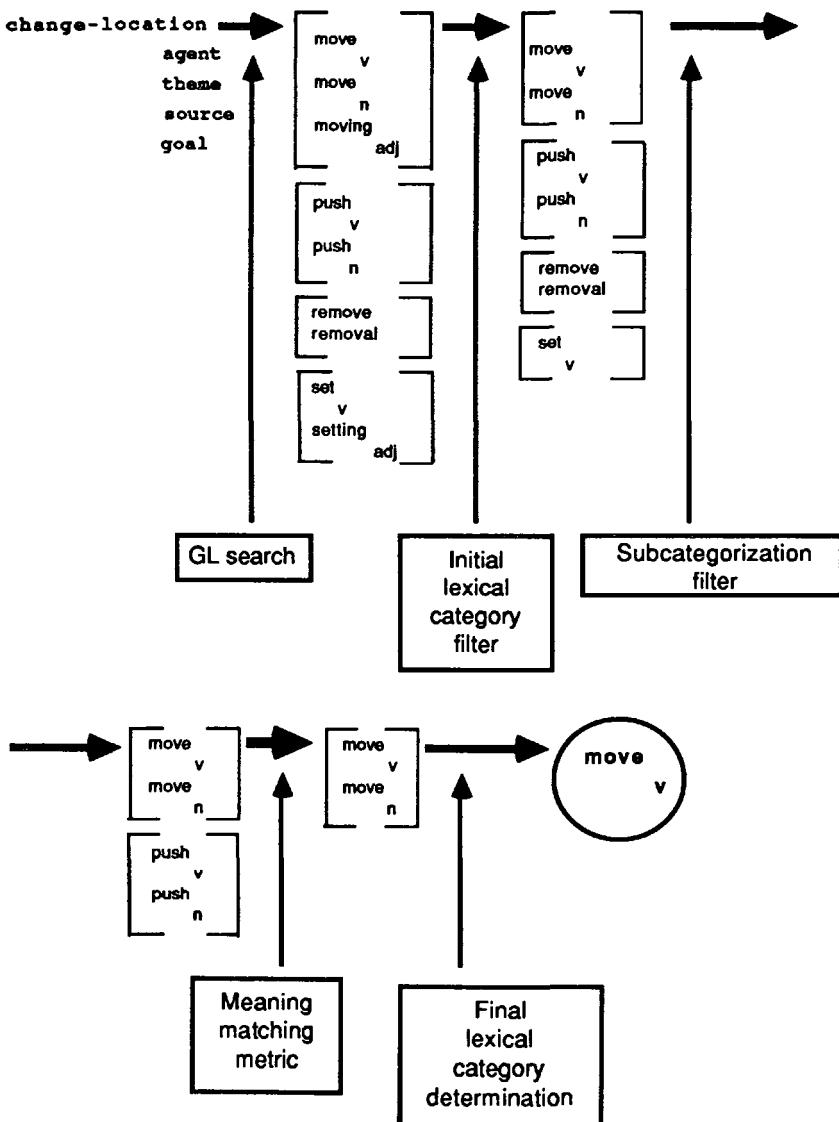


Figure 6. An illustration of the process of lexical selection.

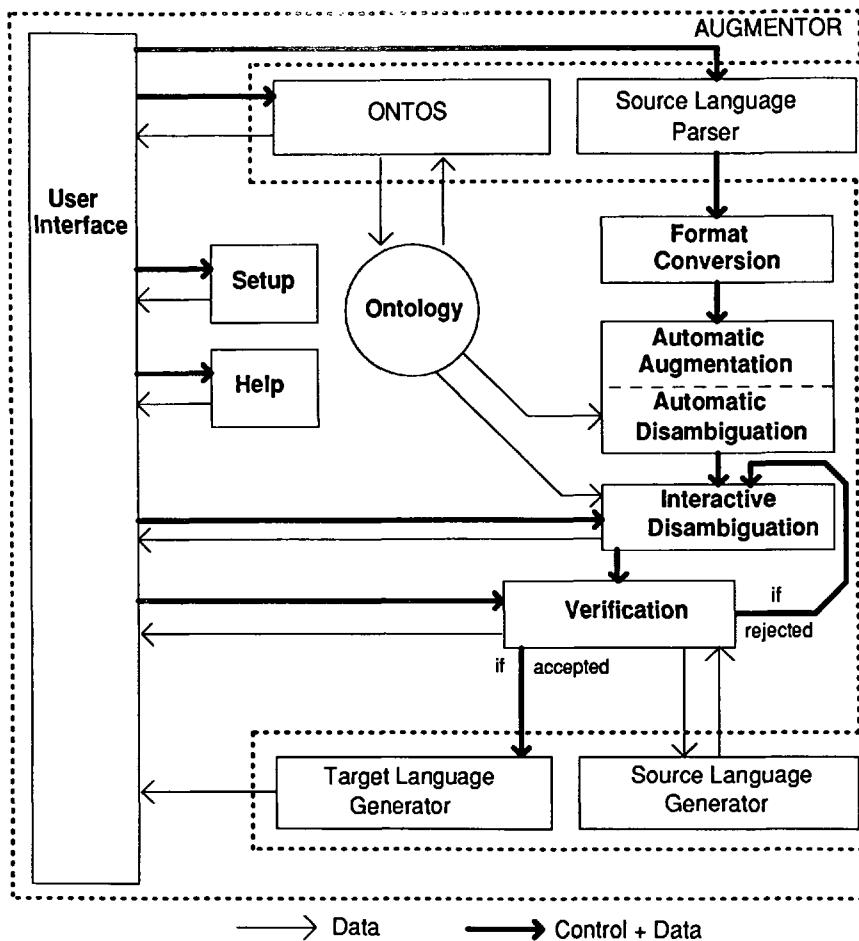


The human-computer interface that would support the interaction of the system and the human user, the interactive editor, has been implemented on a small scale in the MIND machine translation project at RAND Corporation (Kay 1973; the module was called 'disambiguator'). In KBMT-89 the interactive module (known as the 'augmentor') has been implemented on a larger scale.

In fact, the augmentor serves not only as an interface. It is a general filter between analysis and generation. It has an automatic component, which in KBMT-89 deals with referential ambiguity resolution and assignment of non-propositional meanings. The augmentor can also be used for knowledge format modifications that may become necessary if an independently developed analysis module is integrated with the system. The interactive component of the augmentor queries the user about the residual lexical ambiguities, residual problems in attachment of prepositional phrases and subordinate clauses, properties on which nominal modifiers are linked to the heads in noun-noun compounds, etc.

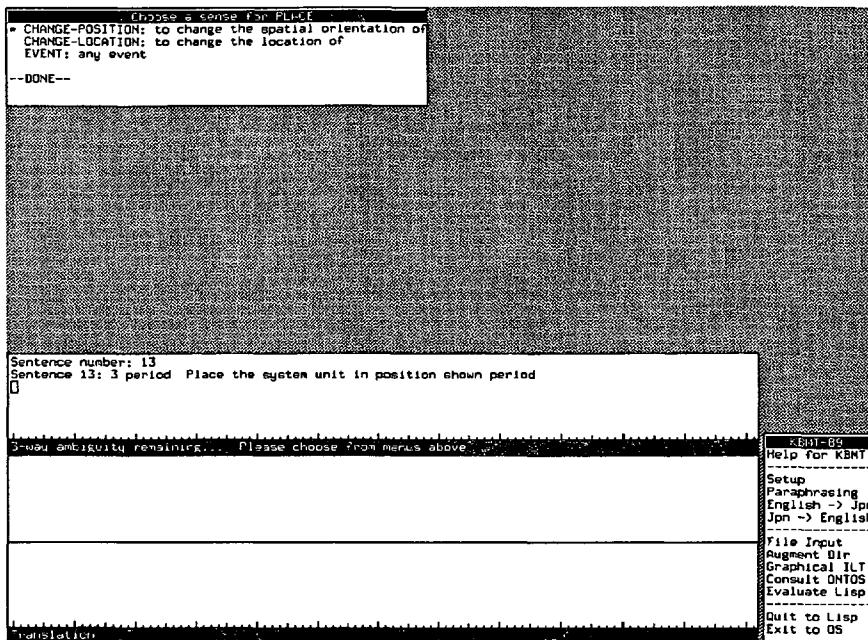
Figure 7 shows the architecture of the KBMT-89 augmentor, while Figure 8 presents a sample augmentor session.

Figure 7. A representation of the Augmentor architecture.



Since the augmentor is integrated in the general user interface of the entire system, Figure 8 also shows the view of the screen during system operation.

Figure 8. Choosing among the remaining candidates.



2.5 MT as an experimental testbed for computational linguistics. In addition to its utility as a machine translation shell, KBMT-89 (or a similar KBMT system) can be very profitably used as a research tool and testbed in computational linguistics and artificial intelligence. To illustrate briefly:

- In its current state the system provides an excellent tool for devising and testing new and more powerful specialized semantic interpretation algorithms, such as, for instance, noun-noun compound understanding or prepositional phrase attachment. With more types of semantic and pragmatic knowledge appearing in the ILT, more specialized 'microtheories'⁴ will be devised and/or incorporated into the process.
- The generation component of KBMT-89 is a very good substrate on which to build more sophisticated natural language generators. In particular, it facilitates the interaction of syntactic, lexical and prosodic

processing and offers a level of reliance on world knowledge that is unusual in most current natural language generators.

- An additional advantage of using KBMT systems as research vehicles is that it is a *comprehensive* system that allows immediate testing of a new component, such as a new parser or a generator, in the context where a 'real' output can be obtained.
- The interface component of a KBMT can serve as a medium for building other interfaces, notably for the purpose of computer-aided instruction and, in particular, for teaching foreign languages. The interface can also be very useful in machine learning systems, especially those studying learning from text or learning by being told, or in systems that investigate hybrid learning processes which involve natural language.
- A comprehensive understanding-and-generation system like KBMT-89 can also be used as a component in a system modelling a cognitive agent--alongside other modules, such as planning and problem solving, perception and action simulation components.
- The ontological and domain knowledge in KBMT-89 can serve as a tool for research in the area of acquisition and maintenance of large knowledge bases. In fact, Ontos is already being used to build domain models in the fields of molecular biology, law, financial transactions and computer software in the framework of projects in the areas of diagnostic expert systems, qualitative process theory and computer-aided instruction. The domain models can also serve as the underlying substrate for a hypertext-type index into a large corpus of human-readable information.
- The computing technology embodied by a KBMT system can be used in other applications. One of the areas in which KBMT-89 can yield immediate practical results is design and development of high-quality translator's workstations. The interaction environment can be extended to include additional types of human-computer interaction. Additional knowledge sources can be connected to the system (for instance, human-readable dictionaries and encyclopedias). And the presence of working analyzer and generator modules will allow the system to suggest acceptable solutions (or informed choices) to the human translator; this is a feature not present in any current translator's workstation.
- Outside of machine translation proper, the technology developed in KBMT-89 is readily usable in applications that require different types of inputs and/or outputs to a natural language processor. Thus, instead of forwarding an intralingua text to the generator, one can pass it on to a special reasoning program that will produce an abstract of the input text, or answer questions based on it, or categorize the

input text into one of a number of taxonomic classes. KBMT-89 (and other knowledge-based machine translation systems) can also be reconfigured for supporting natural language interfaces to database systems. Indeed, if a data manipulation (query) language is substituted for the interlingua, the task of query formulation can become quite similar to that of analyzing a natural language input for translation.

Notes

1. A very good example of what happens when an interlingual system chooses a human-oriented language as the interlingua is the DLT project (Witkam 1987). This project has selected Esperanto as the interlingua and ended up with the necessity of a significant overhaul of the language to make it support the types of processing than an interlingua must support.

2. Such event instances are sometimes called 'episodes'. Their inclusion, together with object instances, into a domain model is a precondition for designing systems that would automatically augment domain models, based on the experience of reading and understanding texts. Such systems can be included as components in future knowledge-based machine translation systems.

3. The KBMT-89 generator is actually a subset of DIOGENES-88, a distributed natural language generator developed at CMU (see Nirenburg et al. 1988). The latter system involves even more knowledge in the lexical selection process (in particular, the knowlege of language-dependent lexical collocations is used; see Nirenburg and Nirenburg 1988 for a description).

4. The concept of microtheory that we use here has been described in greater detail in Nirenburg and Pustejovsky (1988). Provided that a comprehensive, computationally relevant theory of semantic and pragmatic interpretation is not about to be advanced, the best policy for computational linguists in building comprehensive, even though sublanguage-dependent, natural language processing systems is to combine, to the best of their ability, the results offered by partial theories (or, *microtheories*) of particular semantic and pragmatic phenomena (e.g. quantification, reference, thematic structure, discourse cohesion, aspect, time and tense, metaphor, metonymy, etc.). Under this approach, we can use, we can operate without an integrated semantic and pragmatic theory, once we make sure that the findings are represented in a uniform way and introduce a computational control structure that will allow a high degree of autonomy to the component microtheory-based modules, while at the same time maintaining interdependence of microtheory-dependent choices.

References

- Hutchins, W. 1986. Machine translation: Past, present, future. Chichester, U.K.: Ellis Horwood Ltd.
- Kay, M. 1973. The MIND system. In: R. Rustin, ed., Natural language processing. New York: Algorithmics Press.

- Nirenburg, S., and J. Carbonell. 1987. Integrating discourse pragmatics and propositional knowledge in multilingual natural language processing, vol. 2.2. Montréal: The Cognitive Science Society.
- Nirenburg, S., I. Monarch, T. Kaufmann, I. Nirenburg, and J. Carbonell. 1988a. Acquisition of very large knowledge bases. Technical report, Center for Machine Translation, Carnegie Mellon University.
- Nirenburg, S., R. McCardell, E. Nyberg, P. Werner, S. Huffman, E. Kenschaft, and I. Nirenburg. 1988b. DIOGENES-88. Technical Report, Center for Machine Translation, Carnegie Mellon University.
- Uchida, H. 1987. ATLAS: Fujitsu machine translation system. Proceedings of MT Summit, Tokyo, September.
- Witkan, T. 1987. Interlingual MT—an industrial initiative. Proceedings of MT Summit, Tokyo, September.
- Zarechnak, M. 1979. The history of machine translation. In: B. Henisz-Dostert, R. Ross Macdonald, and M. Zarechnak, eds., Machine translation. Trends in linguistics: Studies and monographs, vol. 11. The Hague: Mouton.

Culture in the language class: Videos to bridge the gap

Ricardo M. Paiva
Georgetown University

In the last decades, the focus in the teaching of foreign languages has been increasingly oriented towards fuller participation of the students in activities aimed at improving their ability to communicate in the target language. The Grammar-Translation Method has given way to new approaches based on linguistic, psychological and sociological aspects of language and second language acquisition. One after the other, new methodological approaches known as Audio-lingual, Communicative Competence, Silent Way, Counseling-learning/Community Language, Suggestopedia, and Natural Approach have taken their turn in implementing innovative curricula. Emphasizing different aspects in the process of second language acquisition as a basis for their assertions and proposals, these methodologies appear at times to be in conflict with each other. In reality, all of them have effectively contributed in one way or another to a better understanding of various aspects pertinent to second language acquisition. The experienced language instructor will find in each of them those aspects that definitely work in the language learning process and some which do not with the same degree of satisfaction. One may go even further in saying that any methodology works well depending on the teacher and on his audience. However, the successful teacher, regardless of the methodology applied in the textbook used, will be the one who emphasizes communication through a motivating, creative, and efficient system of interaction, from the first encounter with the language. Once the learning objectives of the students in that particular class are established and the constraints imposed on that specific language environment are defined, the well-advised teacher will design the curriculum around the most appropriate interactive activities. This successful curriculum will then be the one which fits the needs, capacity and interest of the students and which at the same time allows the instructor to feel comfortable with classroom interactions. While expressing his personality in the selection of classroom activities and the design of the course, the instructor is also being true to his own preferences and individuality. What we have here is customized language instruction and language learning apparatus operating through a system of interaction between the instructor and the students, among the students, and between the student and the instructional material.

Interaction presupposes the existence of a message which is both expressed and received in a communicative process which involves exchange of ideas and conveying of intentions, constantly negotiating meaning in a given

context. Furthermore, as Wilga Rivers states, 'the students achieve facility in using a language when their attention is focused on conveying and receiving authentic messages (that is, messages which contain information of interest to speaker and listener in a situation of importance to both). This is interaction' (Rivers 1987: 4). The classroom atmosphere should encourage the students to contribute their personal experience and acquired knowledge always expressing their individuality. The students will feel greatly satisfied when they express what is significant to them instead of repeating impersonal formulas imposed by the textbook. In a more advanced stage, individuality can be meaningfully expressed in making a judgment, formulating a concept, establishing parameters, analyzing, comparing, and drawing up theories.

Authentic messages, with the power they carry to maintain continued interest in interactive communication, originate from authentic situations and materials. An infinite number of situations can be listed as of importance to listeners and speakers, based on their personal experience, interest or knowledge. Typical of these are the problem-solving situations and role plays. Authentic materials can be found in a multitude of sources, such as advertisements, menus, songs, film listings, shopping guides, newspaper or magazine articles, and so forth. Besides the immediate function of providing stimuli for interaction, the authentic materials are a natural way to introduce the student to the culture of the country and to the real language, i.e., the one which comes with a cultural environment.

However, in a regular classroom situation, authentic materials presented in printed form demand the participation of the instructor to enliven them and give them a greater cultural dimension. Also, the teacher's discourse, limited as it is to the conditions of the environment, becomes the only model the students have for imitation. They also learn to associate the teacher's style and manner of expression in the target language with that of all native speakers. When confronted with these native speakers, the students may experience a degree of difficulty in decoding the message received because they may not be familiar with many of the aspects of its code, including differences in pronunciation, intonation, voice quality, and verbal and nonverbal means of expression. The textbook and supplementary materials provide the student with the necessary elements for communication, but success in interactive activities always depends on the capacity of the group to imagine scenarios, create problems and propose solutions to them, design a situation and function within its boundaries. The intensity and effectiveness of this interaction, determining the degree of the learner's communicative competence, will be largely limited to the linguistic expression, devoid of much of the real cultural environment which exists only in the country of the target language.

The introduction of videos in the classroom brings an interlude to these activities. Videos restore the illusion of reality in visual form, entertain, feed the imagination and renew interest in the subject matter. They create scenarios in which students actually see native speakers interacting among themselves, surrounded by cultural trappings that can be shown only through visual images. When the students either imitate or recreate those scenarios or imagine new ones, the original cultural clues provided by the visual images should always work as a point of reference.

Videos are not the perfect substitute for real situations, but their use makes it possible to create authentic visual and linguistic materials to narrow the gap between language and culture. In order to achieve this objective, a series of video capsules should be carefully designed to segment the cultural aspects. These capsules give the student direct access to the culture of the country being studied. They become items to be observed, understood, interpreted, talked about, imitated, and compared with similar situations familiar to the students in their native culture or in others that they know. Instead of merely hearing from the teacher about cultural aspects, which would turn the language class into a lecture class, the students are directed to discover them, and to communicate their findings in meaningful interactive activities. Although imagination still plays an important role in these activities, the culture capsules are an intriguing point of reference. The cultural background of the students, their personal experience, knowledge, maturity, and capacity to make sound judgments all come forward to express, in the target language, a point of view in regard to the stimulus presented in the video.

With videos, the movement is always from comprehension to production, as also happens in real life situations. Krashen and Terrell, among other scholars, maintain that all the student needs is plenty of comprehensible input and 'the ability to speak (or write) fluently in a second language will come on its own with time' (Krashen and Terrell 1983:32). Listening to authentic video segments exposes the students to the language of different social classes, to individual and regional accents, pronunciation, intonation, body language and mannerisms. This real language is different from the one he reads in texts and dialogues because it is produced in an authentic cultural context and is performed by native speakers of all ages and physical types.

In *Travessia*, a Portuguese language textbook used in conjunction with a video program, the information the student receives in regard to real language usage is greatly enhanced by showing native speakers from all regions of Brazil, from different social classes and levels of education communicating in a variety of situations. The language presented in the textbook falls into two broad categories: the language of the author, which is in conformity with its best use as far as accuracy and correctness are concerned, and the language in the dialogues, with a variety of discourse styles and register. In both cases, the goal is authenticity of expression. This authenticity has a much greater dimension in the video because, here, the criteria for the choice of phraseology, expressions, and conversational styles are the high frequency of the items and native naturalness. A unique contribution of the video is in the way it presents a variety of pronunciation and intonation. For instance, Portuguese words such as *está* and *você* are pronounced *tá* and *cê* in informal speech. While it is not advisable to show these short forms in the textbook to avoid confusion in spelling, it is quite appropriate to introduce them to the students in an actual conversation involving native speakers because this is the kind of language they will be exposed to in Brazil. The same situation exists in a very large number of features which differentiate the spoken from the written language in Portuguese. Traditional Brazilian Portuguese textbooks prefer to maintain a safe position by avoiding a display of the real spoken language. Instead, the emphasis is on traditional grammatical correctness,

regardless of its high or low frequency in the spoken language. The common argument to justify this is that the student's acquisition of 'bad language habits' will reflect on the written language and that a distinction between the two forms of the language will be difficult to master. The use of videos can greatly help in solving this problem. It allows the textbook to present the most generally accepted forms of good standard dialect and at the same time introduces dramatizations, in video format, with authentic forms of discourse. With time, the student will identify the level of language spoken in a variety of social situations by native speakers of different cultural levels.

The model for the presentation of different styles in language communication, as shown in *Travessia*, consists of segmenting culture features into capsules. They may be short documentaries on a wide variety of subjects, demonstrations in the use of survival skills, short scenes from Brazilian TV miniseries or soap operas, skits, interviews, songs, dances, folklore, dramatization of legends, popular festivities, national celebrations, and sociological aspects such as religion, or popular beliefs. Each of them is designed to serve a specific purpose. For instance, skits have a controlled vocabulary and provide practice for some grammatical features, without losing linguistic authenticity. They are humorous, artistic, and could be performed by the students in their original form or paraphrased. Interviews with old and recent immigrants to Brazil, with movie stars, singers, cowboys in their work place, pilgrims on their way to a sacred site, and others in a variety of social positions are also very effective devices to bring to the classroom personal experiences and points of view concerning Brazil and Brazilian society. They also serve as models for interviews to be performed by the students among themselves or with native speakers in the community. Songs are carefully selected for their linguistic and cultural content, as well as artistic appeal.

What characterizes all of these cultural capsules is their undeniable mark of being authentically Brazilian. Some of the most distinctive characteristics of Brazilians are their sense of humor, their capacity to laugh at themselves, their warmth, the vibrant way in which they express themselves, their peculiar way of finding solutions to problems, their individuality, their enchantment with Brazil's natural beauty, their keen awareness of regional differences, and the passion for their music. Conversely, those human characteristics which are not applicable to Brazilians are also noticed by default as they are not apparent on the screen. The students will intuitively identify these non-characteristics through a process of nonidentification with their own cultural profile and value system. While some Brazilian characteristics such as customs, preferences, practices, personality traits, and mannerisms may be the topic of a particular capsule, they can be generally observed as secondary components in many scenes. Since an image speaks more than a thousand words, the students will register their observations and will attempt to retrieve them when required to support an opinion, make an assertion or establish a comparison.

In more advanced stages, the students are introduced to the geography of the five regions of the country, their history, economic development, sociological features, architecture, folklore, artistic expressions such as music and fine arts, peculiarities and curiosities. The texts which accompany these features are brief, but the cultural information from the visual image is quite

extensive, providing a number of scenarios for classroom interaction. Compacted miniseries from Globo TV Network of Brazil are also presented. Divided into segments of 10 minutes each to facilitate their use in the classroom, they are supplemented by readings on cultural aspects, with realia and illustrations to create new scenarios for interactive activities based on the cultural aspects observed and analyzed.

The student will learn that reading a text and 'reading an image' on the screen do not demand the same tools. In the reading of a text, there is an interaction of the student with the printed word, which will come to life through the intellect and imagination of the reader. On the screen, words are only a small part of the cinematic language and narrative structure. It is the camera which stores in the mind of the viewer an enormous amount of non-verbal information and provides the material for a large variety of associations.

After familiarizing themselves with the language and situations of a given video segment, the students will engage in activities such as problem solving, role play, imitation, debate of a topic, and so forth, in an attempt to integrate the acquired perception of the culture into their discourse. Above all, they express meaningful messages, thus acquiring a feeling for the language and its culture.

One of the arguments against the inclusion of culture in a language course is that the material presented may reinforce existing clichés or create stereotypes and that the students may be offended by cultural aspects contrary to their own value systems. While these assumptions might be true, two points should be made here: First, ignoring clichés and stereotypes of the culture in question will not make them go away. It is better to present the cultural facts, talk about them and analyze them in their real context. Furthermore, the inclusion of a large number of diversified aspects of the culture will in their own right contradict the clichés and modify the wrong perceptions the students might have. In this way, misunderstandings which are responsible for the existence of clichés have a greater chance of fading out. Second, we must ask ourselves, 'Should we avoid presenting those cultural aspects which are relevant and fundamental to the people of the country being studied, but contrary to the personal preferences and value system of our students?' In doing so, we would, on the one hand, ignore cultural authenticity--linguistic or sociological--while on the other hand accepting cultural imperialism. An example of this exists in the portrayal of gender roles and in male/female relationships. The visibility of this situation is unavoidable and deserves a discussion in the classroom. The role of the teacher here should be that of the educator who tries to guide his students beyond the boundaries of their native culture and into the realm of other cultures. Students should be taught that value systems which are foreign to their own should be carefully dealt with if knowledge of the language and culture they are studying is to be validated by authenticity. Furthermore, in the study of language and culture there is an excellent opportunity for the students to develop their human potential through the understanding of themselves and of the global environment. As an epigraph to *Travessia* (which means 'crossing' or 'going through'), I wrote, 'To learn a language and study a culture is to cross (go through) a world of immense horizons and, in the process, find yourself, as you discover the other.'

If interactive activities in the classroom are an important component in the second language acquisition process, then the message being exchanged must be meaningful in order to maintain a continued level of interest. Meaningful activities are motivated by authentic printed materials and situations created to simulate those which exist in the country of the target language. Consequently, the teaching of aspects of culture is indispensable if the student is to produce authentic language in a meaningful exchange of messages. A carefully planned video program is an asset in the language class because it brings, in dynamic images, a world of information which can be explored in a variety of ways and exploited to familiarize the student with the authentic ways in which the society of the target language operates.

References

- Rivers, Wilga M., et al. 1988. *Interactive language teaching*, 2nd ed. New York: Cambridge University Press.
- Krashen, Steven D., and T. D. Terrell. 1983. *The natural approach: Language acquisition in the classroom*. Oxford: Pergamon Press and Hayward, Calif.: Alemany Press.
- Tolman, Jon M., Ricardo M. Paiva, John B. Jensen, and Nivea P. Parsons. 1988. *Travessia*. Washington, D.C.: Georgetown University Press. The video component of the *Travessia* project was produced with a grant from the National Endowment for the Humanities in conjunction with Globo TV Network and the Roberto Marinho Foundation, both of Brazil.

Using unification grammars for analysis and synthesis

Margaret King
ISSCO and ETI
Université de Génève

Foreword. Despite the single authorship of the paper, the work reported here is that of quite a large group of people; since the author's rôle was relatively limited, it seems only appropriate to open this paper with a list of them. Rod Johnson was primarily responsible for the overall shape of the system, and he, Mike Rosner, Dominique Petitpierre and John Carroll are responsible for the development of the software tools described. The linguistic descriptions of German and of French are the work of Susan Warwick, C.J. Rupp, Graham Russell, and Thérèse Torris. Supplementary work on dictionary coding for German, French, and Italian has been contributed by Riccardo Boschetti, Kirsten Falkedal, Nora Nadjarian, Pascale Dhoop, Sandra Manzi and Lucia Tovena.

Despite the impression of disparate areas of responsibility suggested by the above, it should be emphasized that one of the main principles underlying this work is that, at the design level, there should be a strong mutual influence between software specification and the expressive needs of linguists. Hence, we have adopted a philosophy of rapid prototyping whereby each version of the system is prototyped, is used in linguistic work, and the experience thus gained influences the following version.

This does not mean, of course, that the linguist should have to worry about computational considerations while he is actually constructing a linguistic description: on the contrary, we take as a further principle that the language used for writing linguistic descriptions should be natural for use by a linguist, that is, it should resemble closely the kinds of formalism with which he is familiar from work in theoretical linguistics, and that he should not have to concern himself with how the software of the system applies the linguistic description in order to fulfill a particular task. In particular, he should not have to concern himself with procedural questions, such as the order in which rules are applied or in which computational procedures are executed.

Some background. Before going into any more detail about the prototype we are currently working on at ISSCO, it will be useful to explain a little about the background of the work. Some eighteen months ago, in the autumn of 1987, a Swiss association (Swisstra) concerned with establishing an expertise in machine translation within Switzerland, gave us a mandate to investigate methodologies for the evaluation of machine translation systems. Influenced partly by work done at Hewlett Packard on establishing test suites for natural

language analyzers working on English, partly by techniques used in software engineering for proving the correctness of compilers, we set out to investigate the feasibility of setting up benchmark tests to be used in determining the coverage of a machine translation system. In order to experiment with tests of this type, we needed access to a machine translation system whose functioning we understood well; on the grounds that those who construct something best understand how what has been constructed works, we then decided to construct a small research prototype ourselves.

Since the prototype (known by the rather inelegant name of UD) is part of the work on evaluation, and is not intended as a commercial prototype, we have been able to allow ourselves to make use of relatively untried technology. This primarily manifests itself through the use of unification grammars both for analysis and for synthesis, in an attempt to ensure that the same linguistic description can be used both for analysis and for synthesis, in the influence of recent work in theoretical linguistics on the way the linguistic descriptions are written (much more weight is given to the lexicon than is traditionally the case in machine translation systems, for example), and in the reflection of work in situation semantics to be found in some of the linguistic representations. It would be impossible to discuss all of this here, where I shall mostly concentrate on the language used for linguistic descriptions. The interested reader is referred to Johnson and Rosner 1989, Rupp 1989, and Carroll et al. forthcoming, for discussion of some of the other questions.

Unification-based approaches to grammar. A basic property of unification-based formalisms is the use of feature structures as their informational domain. Different formalisms differ in the way feature structures are defined and represented: cf., for example, the 'terms' of DCGs, the directed acyclic graphs of PATR-II, GPSG's feature bundles, FUG's 'functional structures', and LFG's 'f-structures'. In UD, a feature structure is a set of attribute-value pairs, where a value may be atomic, as in:

```
| number = singular |
```

may be a path, as in

```
| agreement: number = singular |
```

or may be complex. A complex value may consist of a set of attributes and their associated values, as in:

```
| agreement: number = singular
  person = 3 |
```

or may be a reentrant feature structure, i.e. a structure where two or more attributes share the same value, as in:

```
| head: category = vp
  agreement: <#1> | number = singular
    person = 3 | |
```



```
| subject: | category = np | |
```

| agreement => #1 |

Feature structures are combined by unification. The definition of unification used by UD is that two feature structures may be unified providing that they are not mutually inconsistent, the result being a third feature structure which contains all the information contained in each of the two original feature structures. Thus, given three feature structures:

A = | category = adjective
agreement: case = dative |

B = | category = noun
agreement: case = dative
number = singular
gender = masculine |

C = | agreement: case = dative
number = singular
gender = masculine |

then

A |_| C = | category = adjective
agreement: case = dative
number = singular
gender = masculine |

B |_| C = | category = noun
agreement: case = dative
number = singular
gender = masculine |

A |_| B = failure

The grammar rules of UD take the form of classical phrase-structure rules annotated by a set of equations which express the constraints to be satisfied by the feature structures referred to in the two sides of the rule. Thus, the following is a conventional $s \rightarrow np\ vp$ rule, which additionally says that the np should be interpreted as the subject of the vp , and that there is agreement between the subject and the verb:

$s \rightarrow np\ vp$

< s head > = < vp head >
< s head subject > = np head >
< np agreement > = < vp agreement >

The appendix contains a complete grammar for a fragment of English, adapted by Johnson (1988) from an example given in Schieber (1986).

Special characteristics of UD. So far, what has been described is a fairly conventional unification-based grammar. UD offers a number of extensions, including the use of lists and Prolog-like terms as data types. But perhaps the most interesting extension is the possibility to use 'relational abstractions'. Essentially, a relational abstraction allows the linguist to express a generalization once, and to refer to it as often as he wishes in the constraint equations. (A simple concrete example would be to express agreement constraints, instead of doing so explicitly as in the rule given above.) The power of this facility is most obvious, perhaps, at the level of the lexicon, where it allows the work of defining lexical information to be separated from the laborious work of coding the actual entries: the linguist constructing the grammar defines the abstractions, the dictionary coder needs only to know the name of the abstraction and how to use it. Thus, a lexical entry for the German verb *sehen* will appear as:

```
sehen * v/adj
{ - prefix } !Pref (none)
!Nonrefl !Loctype ([] )
!Subcat (np (nom), np (acc), vp (bse), sor )
```

where all the elements preceded by ! refer to relational abstractions. The same entry, when fully expanded, gives a great deal of detailed linguistic information:

```
cat = v
form = sehen
gin = #57
gout = #57
head: morph: sep = none
      sem: desc: arg: 1 = #52
            2 = #41
      loc: cond = [ overlap ( <31> type = loc
                                val = loc.12
                                <34> type = loc
                                val = loc.11)
      overlap ( #31
                <28> type = loc ) ]
      ind => #34
      pol = 1
      rel = sehen
disc: loc: ind => #28
ref: loc: ind => #31
syn: infl: agr: <19> num = pl
      pers = 1/3
tns: fin = pres
     fut = no
     perf = no
mood: = subjunc
perf-aux = none
pref = none
refl = none
subj: <45> sem: desc -> #52
      syn: infl: agr => #19
```

```

        case = nom
        nform = norm
    vform = fin
    voice = active
    vtype = norm
null = no
subcat = [ <22> cat = np
            head => #45
            subcat = []
<23> cat = np
            head: sem: desc -> #41
            syn: infl: case = acc
            nform = norm ]

```

More detail about the dictionaries used with UD can be found in Warwick (1986).

Current status. At the moment, substantial linguistic descriptions for German and French have been constructed and tested for analysis. (The German lexicon, which is the larger, contains around 2,400 items, of which slightly less than half are verb entries). Our main current preoccupation is synthesis: a first version of a synthesizer which uses the same linguistic descriptions as as analysis has recently been completed, and is now being used to investigate what constraints are necessary for synthesis to run satisfactorily. Transfer will be added, we hope, later this year, in the light of extensive experimentation with the parallel linguistic descriptions.

Appendix: A simple UD grammar.

Taken from Johnson (1988), adapted from Shieber (1988).

Declare

```
Category = cat
```

Grammar

Rule sentence formation

```

$ -> NP VP
< * cat> = s
< NP cat> = np
< VP cat> = vp
< * head form> = finite
< VP subcat> = [NP]
< * head> = < Vp head>

```

Rule trivial verb phrase

```
VP -> V
```

```

< VP cat> = vp
< V cat> = v
< VP subcat> = < V subcat>
< VP head> = < V head>

```

Rule complements

```

VP -> Fun Arg
< VP cat> = < Fun cat> = vp
< Arg cat> = np/vp
< Fun subcat> = [ Arg | < Vp subcat> ]
< VP head> = < Fun head>

```

Lexicon root

```

sleeps  v   < * head form> = finite
          < * subcat> = [ NP ]
          < NP cat> = np
          !Agree (NP, 3, singular)
          < * head sem pred> = sleep
          !Arg ( 1, NP)

sleep   v   < * head form> = finite
          < * subcat> = [ NP ]
          < NP cat> = np
          !Agree (NP, _, plural)
          < * head sem pred> = sleep
          !Arg ( 1, NP)

sleep   v   < * head form> = nonfinite
          < * subcat> = [ NP ]
          < NP cat> = np
          < * head sem pred> = sleep
          !Arg ( 1, NP)

storms  v   < * head form> = finite
          < * subcat > = [ Obj, Subj ]
          < Obj cat> = < Subj cat> = np
          !Agree ( Subj, 3, singular)
          < * head sem pred> = storm
          !Arg ( 1, Subj)
          !Arg ( 2, Obj)

storm    v   < * head form> = finite
          < * subcat > = [ Obj, Subj ]
          < Obj cat> = < Subj cat> = np
          !Agree ( Subj, _, plural)
          < * head sem pred> = storm
          !Arg ( 1, Subj)
          !Arg ( 2, Obj)

storm    v   < * head form> = nonfinite
          < * subcat > = [ Obj, Subj ]
          < Obj cat> = < Subj cat> = np
          < * head sem pred> = storm

```

```

!Arg ( 1, Subj)
!Arg ( 2, Obj)

stormed v   < * head form> = pastpart
             < * subcat > = [ Obj, Subj ]
             < Obj cat> = < Subj cat> = np
             < * head sem pred> = storm
             !Arg ( 1, Subj)
             !Arg ( 2, Obj)

has      v   < * head form> = finite
             < * subcat > = [ Vcomp, Subj ]
             < Vcomp cat> = vp
             < Vcomp head form> = pastpart
             < Vcomp subcat> = [ Subj ]
             < Subj cat> = np
             !Agree ( Subj, 3, singular)
             < * head sem pred> = perfective
             !Arg ( 1, Vcomp)

have     v   < * head form> = finite
             < * subcat > = [ Vcomp, Subj ]
             < Vcomp cat> = vp
             < Vcomp head form> = pastpart
             < Vcomp subcat> = [ Subj ]
             < Subj cat> = np
             !Agree ( Subj, _, plural )
             < * head sem pred> = perfective
             !Arg ( 1, Vcomp)

persuades v   < * head form> = finite
             < * subcat > = [ Obj, Vcomp, Subj ]
             < Obj cat> = < Subj cat> = np
             < Vcomp cat> = vp
             < Vcomp head form> = infinitive
             < Vcomp subcat> = [ Obj ]
             !Agree ( Subj, 3, singular )
             < * head sem pred> = persuade
             !Arg ( 1, Subj )
             !Arg ( 2, Obj )
             !Arg ( 3, Vcomp)

to       v   < * head form> = infinitive
             < * subcat > = [ VP, NP]
             < NP cat> = np
             < VP cat> = vp
             < VP head form> = nonfinite
             < VP subcat > = [ NP ]
             < * head sem> = < VP head sem>

arthur   np  !Agree ( *, 3, singular)
             < * head sem> = arthur

cornwall np  !Agree ( *, 3, singular)
             < * head sem> = cornwall

knights  np  !Agree ( *, 3, plural)

```

```

< * head sem> = knights

# Define relations

Agree ( X,
        < X head agreement person >,
        < X head agreement number> )

Arg ( N, X )< * head sem arg N > = < X head sem >

# Lookup root < * cat > = v/np

# Restrict Cat Form
    < * cat> = Cat
    < * head form> = Form

```

References

- Carroll, J., G.J. Russell, and S. Warwick. (forthcoming) Adapting unification grammars for generation. ISSCO WP56.
- Johnson, R. 1988. Le traitement automatique de la langue. In: Workshop on Artificial Intelligence and Robotics, Sion.
- Johnson, R., and M. Rosner. 1989. A rich environment for experimentation with unification grammars. ACL-Europe.
- Rupp, C. J. 1989. Situation semantics and machine translation. ACL-Europe.
- Shieber, S. M. 1986. An introduction to unification based approaches to grammar. Chicago: University of Chicago Press.
- Warwick, S. 1986. Building up a lexicon for NLP applications. LDV-Forum, 5.4, Sept.

Testing and technology in Germany revisited: What is left? What can be hoped for?

Reinhold Freudenstein

Phillips-Universität, Marburg

I still remember quite well when it all began thirty years ago. At that time, foreign language teaching in Germany was very traditional both at school and university levels. After the Second World War, General Clay--at that time responsible for the introduction of a new, democratic German state--is supposed to have said about German higher education: 'Germany has the best university system in the world--for the nineteenth century.' This still was true in the middle of the fifties. Foreign language teaching was organized on the basis of grammar and translation; it was primarily regarded as a contribution towards general education by making it possible to study foreign literature as examples of great cultural achievements. And then, in 1957, the Sputnik shock not only had an effect on the United States but on Europe as well. In America, the National Defense Education Act made the study of foreign languages a matter of national interest; language studies became as important as the natural sciences, and this soon led to the development of a concept of language teaching and learning as a scientific discipline. In this context, the audiolingual approach was born--the successful attempt to integrate linguistics and language teaching in a coherent, psycholinguistic theory. It was here in Georgetown at the School of Languages and Linguistics that Robert Lado started what one could call the modern period of the language sciences. This new, scientific approach to the teaching and learning of foreign and second languages was accepted worldwide. It was particularly welcomed in Germany where the historical experience had led to great doubts about the value of humanistic education. Furthermore, it was felt that traditional language teaching did not prepare students at all for the practical application of their knowledge of a language outside the classroom. Hope for entirely new ways and promising activities was focused on two major fields: education technology and testing. In this paper I would like to give a short report on the way in which we received and adapted the new ideas in language sciences at that time in Germany. I will also report about today's situation and, using this as a basis, it will be possible to comment on future developments. I have subdivided my report in two parts; I will speak about technology and testing in separate chapters because both fields of activity are of a different nature. In addition, they have developed independently of each other and have brought about quite different results.

1 Educational technology. I would like to speak about educational technology first because the impact of technology on the educational system

and particularly on language teaching in the late fifties was much stronger than that of any other field related to education. It was felt on all levels, from administrative policy-making to each and every classroom. Many educators thought technology would entirely change education in the times to come. Today it is the area of technology again that dominates professional discussion; again educators hope for radical changes, and this time technology has come to us in the form of the computer. This makes it particularly interesting to look back to the time when educational technology was born.

1.1 The past. Thirty years ago, educational technology in the language classroom was concentrated on the language laboratory and the new forms of exercises that went along with it. In Germany, the idea of teaching foreign languages with the help of language laboratories became known at a time when in the United States there were already first doubts about their usefulness. But the revolutionary idea introduced with the language laboratory in Europe was not the technical equipment only; more fascinating was the new learning objective connected with it: to learn a foreign language as a means of communication for everyday use. In the sixties, in an experimental program the Volkswagen Foundation spent millions of German marks for the installation of language laboratories in German schools. Soon it became necessary for all institutions that taught foreign languages to have the new learning tool available. The claims and hopes connected with this electronic teaching aid made it virtually impossible for any professional opposition to express itself. In the early seventies no school in the Federal Republic of Germany thought it could do without a language laboratory.

1.2 The present. What is the situation today? As in the United States, the magic about the language laboratory has gone. Only two out of about twenty former German language laboratory producers are left. The majority of teachers thinks a language laboratory is superfluous in the language learning process. There are still--as there have always been and will always be--a few specialists who work with it but most of their colleagues have returned to the blackboard-and-textbook approach, and they feel supported by linguistic and psycholinguistic studies that stress the limitations of structuralism and the weaknesses of behaviorism. However, in spite of the fact that language laboratories are no longer used and pattern drills are violently criticized, the area of the language laboratory has left important traces in present-day language teaching. First of all, there is the importance of the oral skills which are generally accepted and which have become an obligatory part of each foreign language curriculum. Another change has taken place in the production of teaching materials. Many teachers in my country who refuse to go to the language laboratory do not know--or do not want to know--that most forms and types of exercises which they apply in their teaching today were developed in connection with just that medium which they now refuse to use.

Personally, I believe that the language laboratory--correctly and wisely used--can still be of enormous help in the language learning process, and I know quite a few colleagues and institutions who support that view, e.g. directors of language centers at German universities or the language

departments of industrial firms. Therefore, I regret that most school teachers are opposed to it. But German teachers have actually always been media-shy or even hostile towards media use. More important is the fact that the context of language teaching has changed under the influence of technology, and this influence has caused a reorientation of methodology that has become more powerful in its effect than regular visits to the language laboratory and the use of its technical equipment. Without the advance of educational technology language teaching would be less communicative, less interesting and less successful.

1.3 The future. A similar development to that of thirty years ago can be observed again now; highly advanced technology is applied to the learning process and advertised at the beginning of an entirely new era in language learning: the computer. Hardware and software producers flood the market with incredible promises. The terminology of their language seems to be directly taken from the time when the language laboratory was first introduced to the educational market. This is no surprise because most of the operations a computer can perform today had been presented before in connection with the technology of programmed learning. Therefore, technology in the language sciences of today is not something entirely new; it simply makes accessible techniques and forms of learning which less advanced technology of the past was not able to provide.

There is one spectacular difference, however, from the time when the language laboratory was first presented. In spite of the many books and articles on computer-assisted language learning, teachers seem not to be really interested in what science and the producers have to offer. Most of them just ignore the new piece of equipment. Thirty years ago teachers simply did not know how to react to the language laboratory; they accepted it in spite of all the reservations they might have had. Today the same teachers or their successors in office have lost confidence in scientific improvements of language teaching. Thus we have a relatively small group of computer specialists highly trained and ready to work out interesting programs that are technically possible but educationally often unwise or unnecessary. No one can doubt that there will be developments which come along with the computer which can indeed have a positive influence on the learning of languages, e.g. individualism or the creation of new types of exercises. But it is my prediction that in the long run the computer in the foreign language classroom will not be regarded as an indispensable medium of instruction, and language teachers will not go with their classes to the computer room of the school--should there be one--as they used to go the language laboratory. This certainly is true for my country, and I don't think that the situation in other European countries is entirely different.

The longer lasting influence of the computer--as in the case of the language laboratory--will become apparent in a more indirect way. I have already mentioned new types of exercises which the computer can produce at high speed while considering individual learning conditions. With the help of a computer the preparation of worksheets has become easier. What we have to watch very carefully, however, are the position and the value of new ideas and techniques in modern language teaching methodology. There is a certain

danger that in connection with the computer, learning materials will be presented which are old-fashioned and outdated in view of modern teaching philosophies. Many vocabulary programs, for instance, bring the grammar-translation method to life again.

The lesson we have learned from the language laboratory and the computer in foreign language teaching and learning is that educational technology cannot radically change existing educational structures or traditional teaching practices. It will simply widen the spectrum and the possibilities of strategies that a teacher has available and makes use of within the context of his personal preferences. A teacher can -- but need not -- work with new machines or new forms of instruction. But independently of the decision individual teachers take for themselves, a change will always take place in the general teaching and learning environment as soon as a new technological medium comes into existence. This fact is the real and genuine progress in our profession. The changes that are brought about in this way are less spectacular than the promises that go along with the invention of new pieces of equipment, but they will be of greater impact, and they will last longer. Most of all, they will influence the teaching styles of more typical teachers in a more comprehensive way than anything else while the concentration on one specific isolated medium like video or the computer can only have a lasting impact on those few and outstanding teachers who are specifically interested.

2 Testing. My report on testing foreign languages in Germany is simple and short because there is not very much that could be said. There is, of course, a long tradition of subjective evaluation in the educational system of my country. Objective testing, however, has found no place in the school system up to now in spite of many efforts by individuals and institutions that have tried to convince the profession of the advantages of testing.

2.1 The past. 'Testing' is a word which was not known in foreign language pedagogy until Robert Lado's publication *Language Testing* came out (Lado 1961). In Germany, the first article on testing foreign languages was published in a professional journal in 1967; it was not written by a language specialist but by a psychologist who was personally interested in language learning and in introducing objective testing in the foreign language classroom (Ingenkamp 1967). Up to that time we simply spoke of 'Klassenarbeiten'--written documents produced in class--and they consisted of traditional forms like dictation, vocabulary review, or essay writing. Lado's book came out in a German translation ten years after it had first been published in the United States (Lado 1971). When the German version had become available a few teachers began to make use of the advantages of testing. As in the case of modern technology, there were a few specialists who really tried to apply objective testing in their classrooms, and there were even ministries of education that officially allowed them to write two tests instead of one regular, traditional 'Klassenarbeit'. But these teachers were in the minority. What happened was that after a relatively short time of initial contact with testing, interest waned, and since that time nothing has changed in this regard. Testing simply does not exist in German schools. There was

only one interesting consequence. Teachers like the expression 'to test', and it has become quite regular to speak of a 'test' when dictations or essays are written.

2.2 The present. When you look at the list of German 'Schulleistungstests' (Beltz 1977:28-31)--the equivalent of objective tests--you find fewer than ten in connection with foreign languages, e.g. a placement test for students of English after one year of school instruction, a vocabulary test in French for students after the third year of learning, or a proficiency test for advanced learners of English. They were produced ten years ago, and since then no further research has been done in this field. The old tests were not revised and no new ones were published. Official state examinations are not even partially based on forms of objective testing. What is left is the practice of a few teachers to write informal tests for their students. Some forms of testing have established themselves in workbooks published in connection with regular textbooks, and some computer programs offer selected items of testing strategies like the cloze test. But these are exceptions. Testing has found no place in the German state school system.

There is, however, a more encouraging language testing development in the adult sector. Here, from the very beginning, the German Adult Education System--the 'Volkshochschule'--saw the advantages of objective testing for their language programs, and in 1967 they introduced a nationwide 'Certificate Program' that had already been tried out in Bavaria (International Certificate Conference 1984). With this program a learner of English, French, Italian, or some other foreign language can prove that she or he has mastered the linguistic skills needed in everyday situations. The program consists of a detailed curriculum, a description of language functions, situations, and structures, a word list, and other elements which are necessary to make the system transparent to teachers and learners.

Today, this 'Certificate Program' has become international, at least in the European context. Twelve countries are participating, among them France, Norway, and Hungary. It is not unusual for employers to accept the certificate as a proof of sufficient language knowledge, and even high school students take the examination in order to find out how good they really are in spite of bad marks in school.

2.3 The future. Testing in the true meaning of the word will most certainly further develop outside the schools and universities. The 'Certificate Program' was set up as an integrated unit credit system with several modules. In addition to the certificate examination there is a Stage 1 examination which can be taken after two years of instruction in the adult system (which normally means two hours of instruction per week). There are special examinations for English at Work, English for the Hotel and Catering Staff, French for the Office, and further levels are planned in all certificate languages. Among the new test forms is a grouped test. With its help one hopes to find out as objectively as possible how good a candidate is at speaking a foreign language. Certainly, schools will not--and for various reasons probably cannot--follow this development. Tests are considered impersonal and too mechanistic. School officials fear that creativity might be suppressed when tests are

introduced in all schools, and there are misgivings about the possibility of teaching only what afterwards can be tested.

There is a slight hope, however, that this general tendency will change with the integration of Europe after 1992. A United Europe cannot develop language-wise along the lines of the past. A new language policy must be set in motion, and it will probably be a policy of multilingualism (Freudenstein 1988). Instead of learning one or two languages over a longer period of time (four to seven years at the moment), students might well get the chance to acquire three or four languages in compact course form over a period of two years, so that they can pick up enough to be able to build on it later on. In such a system testing might become a necessary element in the instructional process because then the emphasis is not on humanistic values of language learning but on communicative skills. This is what I hope for. But at the same time I know that many of my colleagues hope this would never come true.

References

- Beltz Test Gesellschaft. 1977/78. Tests für die Schule. Fremdsprachen-Leistungstests. Weinheim: Beltz.
- Freudenstein, Reinhold. 1988. Fremdsprachenunterricht für das Jahr 2000. Praxis des neusprachlichen Unterrichts 4.339-48.
- Ingenkamp, Karlheinz. 1967. Die Feststellung der Fremdsprachen-Leistung und -Eignung durch Lehrerurteile und Tests. Praxis des neusprachlichen Unterrichts 3. 277-86.
- International Certificate Conference. 1984. Certificate in English. Frankfurt: Pädagogische Arbeitsstelle des Deutschen Volkshochschul-Verbandes.
- Lado, Robert. 1961. Language testing. The construction and use of foreign language tests. New York: McGraw-Hill.
- Lado, Robert. 1971. Testen im Sprachunterricht. München: Hueber.

New directions of machine translation

Makoto Nagao
Kyoto University

1 Introduction. There are growing demands for language translation in every country because the exchange of persons, information and commercial products among different countries is increasing year by year. Especially strong demands exist in Japan for translation between Japanese and English in the business world. Human translation is slow and cannot fill the demand. Moreover, the quality of human translation is not necessarily good.

We have been conducting research in Japanese-English machine translation at Kyoto University since 1965, and developed a sentence-parsing program, dictionary handling systems, and so on during 1965-75. Then we developed a machine translation system (MT), TITRAN, which translated titles of research papers in science and engineering. We constructed TITRAN-EJ (English to Japanese), TITRAN-JE (Japanese to English), and TITRAN-JF (Japanese to French) until the end of the seventies. TITRAN-EJ was used by many researchers and received favorable acceptance. It translated *He is a boy* into *Helium is a boy* in Japanese! The system was so much specialized in the scientific field.

Then in 1982 we started a national MT project, which was supported by the Agency of Science and Technology of the Japanese Government, and which aimed at the translation of abstracts of scientific and technical papers. In this four-year project we developed two MT systems called Mu-JE (Mu is an acronym of our system) and Mu-EJ. The systems were quite large, and included 70,000 words and 1500 grammatical rules in each language. The system is now under reconstruction for the purpose of connecting it to a large database of the Japan Information Center of Science and Technology.

This research and development activity gave great encouragement to computer companies in Japan. We transferred technology to these companies, and they started the construction of commercial MT systems. At present there are about ten commercial MT systems (half JE and half EJ) and several others under development.

The merit of introducing MT systems for the translation of documents is gradually being realized in Japanese companies, and it is widely recognized that the systems will have increasing acceptance in society in two or three years. This is because the users have become wise enough to use a system optimally in a given environment by understanding the possibilities and the limitations of the present-day MT systems, and by not having excessive expectations as to the translation quality.

The research results so far in academic circles have been almost completely transferred to industry, and basic research is moving toward

completely new models, which will give new impetus to higher level MT systems in the next century.

It is quite important to clarify different components related to machine translation in order to evaluate the state-of-the-art of present-day MT systems and to consider future directions. So far, MT systems have treated one sentence at a time for translation without any reference to previous sentences in a text. However, there are varieties of interconnecting information among adjacent sentences, essential for the interpretation of the individual sentence. Several linguistic and computational linguistic research projects are going on in this area, and useful results will be included in MT in the near future. This will become a basis for more advanced MT systems. Several different factors related to MT are pointed out here, and some future perspectives are given.

2 Classification of sentential styles for MT. Present-day MT systems cannot accept arbitrary sentential styles. They do not accept literary works. They have a lot of trouble with precisely translating patent documents, contract documents, and so on. The systems have great difficulty in analysing long sentences of 50 or more words. Thus we have to be very careful about sentence categories and different sentential styles. There are many categorizations of input sentences for MT systems. The followings are some typical classifications:

A Kind of text.

(1) Newspaper headlines, titles of articles. In this category of sentences the most important factor is the speed of information dissemination. The delay will be a few hours at most. Low quality of translation may be permissible. The level of quality at which the readers can understand the meaning from past information and environmental context will be allowed.

(2) Technical information, technical news, economic and social news, and so on. The speed of information transfer is most important. The delay will be a few days at most. Readers of the translated materials in categories (1) and (2) will be professionals in the text area, not the general public. Therefore they are accustomed to the quality of MT and can properly interpret awkward translations.

(3) Operation and maintenance manuals for devices and systems. The volume is great, and translation will require a longer time. The translation quality may not be high, but must keep a certain level, because the translated materials will be printed and distributed to many customers of the devices and systems. Postediting will be required for machine translated sentences.

(4) Business letters which are prototypical. They can be translated by identifying customary expressions. Postediting will be required so that recipients can read the letters readily and easily.

(5) Scientific and technological documents and research papers. First rate translation quality will not be required because specialists in these areas will read and understand the contents.

(6) Conference documents, contract documents, law documents, etc. These are to be translated very carefully. Not only professional translators but also specialists in the document area are needed for postediting.

(7) Patent documents. A huge amount of patent documents has been accumulated and the number is increasing daily. There is a severe shortage of translators in this area, and MT is expected. However, the sentences in patent documents are complicated and difficult to interpret even for human translators. MT will be difficult without heavy professional pre- and postediting.

(8) Articles in newspapers and journals, which are not too literary.

(9) Literary works.

(10) Dialogues in a restricted task domain. These include man-machine, man-man; written sentential conversation by on-line typewriters.

(11) Free dialogue (written sentential conversation).

(12) Speeches such as lectures (one person).

(13) Spoken dialogue (two or more persons).

B Classification of sentential styles (1).

(1) Sentences expressing facts only.

(2) Sentences which include time relations, expectations, assumptions, and conditions.

(3) Sentences which include the speaker's intention, mental state, and so on.

(4) Dialogue sentences which presuppose hearer's knowledge.

(5) Sentences which include metaphors, culture-specific expressions.

C Classification of sentential styles (2).

(1) Length of a sentence (in Japanese, less than 20 characters, 20-40 characters, 40-70 characters, 70 or more).

(2) Number of predicates in a sentence.

(3) Kind and number of conjunctive phrases and clauses.

(4) Sublanguage (use of specific styles in a particular domain) and other complex embedded structure.

(5) Broken and fragmentary sentences.

3 Linguistic components for MT. There are many linguistic theories to be considered in the interpretation of sentences and texts. Some of these theories are already incorporated in MT, but others are still at the research level and cannot be brought into MT systems. Generally speaking, linguistic theories provide us with the basic philosophy of language. Actual language data (sentences) are so complex that the theories cannot explain the reasons for many of them. Linguistic theories are usually based on a very small set of linguistic phenomena, and explains only a small part of language phenomena. MT must treat all the varieties of linguistic expressions, and there must be consistency among the theories taken into an MT system. There are many varieties of linguistic systems which might be considered for MT systems.

D Linguistic components.

- (1) Kind of grammar formalism (context-free, annotated context-free grammar, transformational grammar, case grammar, unification grammar, lexical functional grammar, generalized phrase structure grammar, etc.).
- (2) Treatment of linguistic semantics (semantic primitives, thesaurus,...).
- (3) Dictionary content (part of speech, pronunciation, semantic information, case information, modality, volition,...).
- (4) Kind of dictionary words (common word, compound word, terminological word).
- (5) Kind of dictionary (monolingual, bilingual, multilingual, concept dictionary).
- (6) Knowledge representation and its use in MT.
- (7) Ambiguity resolution by contextual information.
- (8) Treatment of anaphora, ellipsis, focus, speaker's intention and paraphrase.

4 Aspects of MT systems. MT systems can be classified and considered from many different points of view and aspects. One well-known classification is: fully automatic MT, human-aided MT, and machine-aided human translation. There are many other aspects, some of which are listed here:

E Translation method.

- (1) Transfer method.
- (2) Pivot method.
- (3) Direct method.
- (4) Paraphrasing and MT.
- (5) Learning function in MT.

F Pair of languages.

- (1) From a specific language to another specific one.
- (2) Bidirectional between two languages.
- (3) Multilingual.

G Human intervention.

- (1) Preediting.
- (2) Postediting.
- (3) Preediting and postediting.
- (4) Interactive help at arbitrary stages.
- (5) No human intervention.

H Analysis of translation errors.

- (1) Ambiguity resolution (one candidate, all possible candidates, several candidates in the sequence of plausibility).

- (2) Errors in the analysis phase (conjunctive noun phrase, conjunctive sentential clause, determination of prepositional modifiers, etc.).
- (3) Errors in the semantic interpretation.
- (4) Errors in the selection of a target language word.
- (5) Errors in the generation of a target language sentence.
- (6) Errors caused by lack of contextual information.
- (7) Vagueness of the original sentence.

5 MT Use and market. Construction of reliable MT systems is difficult at the present stage of computational linguistics technology. Users of MT should have some special skills in order to utilize the systems. They are asked to augment the dictionary contents and sometimes grammar rules to tune the system to their own text translation style. Therefore user companies of MT systems must have professional persons to maintain and improve their MT system. We can categorize user classes in the following way:

I User classes of MT systems.

(1) A few user companies for a system (less than ten). User companies are expected to keep professional persons for their systems, and they must cooperate closely with the manufacturers of their MT systems in order to improve those systems.

(2) A small number of users for a system (less than 100). Users must know which details of their system to improve. Dictionary enhancement and improvement are typical examples. Such users maintain only slight contacts with manufacturers for the improvement of their system.

(3) A reasonable number of users for a system (around 1,000). A system must be compact and reliable enough for users to handle it without any help from the manufacturer of the system. Users must be familiar with the system's weaknesses and must have a variety of skills to avoid them and to insure getting reasonable results.

(4) A large number of users for a system (popularized stage). System must have a learning capability and must be adapted to customer's text domain automatically. A system must be very cheap, and people must be able to use it without any trouble in ordinary homes.

6 Present status and future perspectives. MT systems at present are far from satisfactory as a commercial product. However, there are various possibilities for utilizing them in specialized situations, such as quick stock-market reports, quick survey of newspaper articles, and so forth. The most important thing for users is to recognize the capabilities and incapacities of present-day MT systems, and to have the ability to discover suitable application areas.

One immediate possibility is to find a very restricted small area where stylized sentences and a small vocabulary are used. TAUM-METEO is a good example. Authors must learn to write source sentences carefully so that no postediting is necessary. This type of system can be found in categories A-1, 2, B-1, 2, C-1 (less than 40 characters), D-1, 4, 5. Another application area

is that of translation of operation and maintenance manuals for machinery, electronics and so on. The market in this category is growing rapidly in Japan.

Systems in this category presuppose that the input sentences are short and do not have carefully nuanced implications. Sometimes preediting is required as well as postediting, which is essential. These systems will fall in the categories A-3, 4, B-1, 2, C-1 (less than 40 characters and at most 70 characters), D-1, 2, 3, 4, 5, I-1. There are other applications such as business letter translation for quick reading at receiver side, and technical translation combined with information retrieval and database systems. The Japan Information Center of Science and Technology is now constructing a Japanese-English MT system to be attached to its databases. When it is completed, anybody in the world can send a retrieval request to the Center in English, and get retrieved results in English translation by machine.

Several future directions are conceivable for MT systems. One is to improve the present MT systems by giving more accurate information to grammar rules and dictionaries. Readability of translated materials must be greatly improved. The document categories must be enlarged from technical reports to much wider text areas such as A-4 to A-8.

Another direction is a multilingual MT. The Eurotra project is typical. There is a similar project in Japan, a Japanese Government project for translating technical documents between the Japanese and Chinese, Indonesian, Malaysian, and Thai languages. It began two years ago, will continue another four or five years. More futuristic is a Japanese electronic dictionary project supported by our government, which aims at representing all the word meanings in individual concepts (millions of concepts); via these concepts, the correspondence of word usages in different languages will be achieved. This is called a concept dictionary. The project is now in its third year. Japanese and English are the languages at present. The project will continue another several years.

Dialogue translation is another futuristic research topic. We established a research institute called the Automatic Translation Telephone Research (ATR) Institute three years ago in Kansai Research Park. The aim is to construct a prototype of automatic speech translation between Japanese and English in another twelve years. The institute has a variety of research projects from speech analysis/generation, language analysis/generation, knowledge representation in the task domain of a dialogue, and so on. Linguistic theories such as D-3 to 8 are intensively studied for incorporation into the system. There is a growing number of linguistic research projects for dialogue sentences in the United States and in Japan, and we can expect that this research will succeed.

To improve translation quality we must study more intensively the mechanism for understanding language. Proper translation expressions, including word selection and sentential styles, can be obtained by means of the factors which control the coherence of a text, from the reader's knowledge, and from the factors which control his focus of attention. The old vs. new information distinction influences the text styles of the reader, and so on. Japanese researchers are of the opinion that MT technology, taken in a wide sense, will be a common base-technology for the coming information-based society.

Appendix: Commercial MT systems in Japan.

Manufacture	System name	Language	Sets sold
Fujitsu	ATLAS-1	E-J	170
Fujitsu	ATLAS-2	J-E	95
Sharp	DUET	E-J	
Toshiba	AS-TRANSAC	E-J	
Hitachi	HICATS/EJ	E-J	30
Hitachi	HICATS/JE	J-E	70
Sanyo	SWP-7800	J-E	500
Oki-Osaka Gas	PANSEE	J-E	150
NEC	PIVOT	J-E/E-J	30

From: *Nikkan Kougyou Newspaper* (1989.1.30)

References

- Nagao, M., and J. Tsujii. 1976. Analysis of Japanese sentences by using semantic and contextual information. American Journal of Computational Linguistics, microfiche 41.1.
- Nagao, M., J. Tsujii, and J. Nakamura. 1985. The Japanese Government project for machine translation. Computational Linguistics 11.2-3.
- Nagao, M., J. Tsujii, and J. Nakamura. 1986. Machine translation from Japanese to English. Proceedings of the IEEE 74.7:7.
- Nagao, M., J. Tsujii, K. Mitamura, T. Hirakawa, and M. Kume. 1980. A machine translation system from Japanese into English: Another perspective of MT systems. COLING 80.10.
- Nagao, M., J. Tsujii, Y. Ueda, and M. Takiyama. 1980. An attempt to computerized dictionary data bases. COLING 80.10.
- Nakamura, J., J. Tsujii, and M. Nagao. 1985. Grammar writing system (GRADE) of Mu-machine translation project and its characteristics. Journal of Information Processing 8. 2:10.

Machine translation: Achievements, problems, promise

Winfred P. Lehmann

The University of Texas at Austin

1 Achievements. Among its many achievements, the demonstration by the School of Languages and Linguistics that human language can be manipulated by computer is unique. Not surprisingly, the interest in such a possibility, and the demonstration as well, were concerned with a practical benefit, translation. As in many developments, others had concerned themselves with the possibility earlier. Locke has given a clear and precise account of early history, including the patent application filed by Smirnov-Troyanski of Moscow in 1933 (1985:129-32; see also Hutchins 1986:21-24). And Warren Weaver's memorandum of 1949 entitled 'Translation' must be credited for its suggestion. But these efforts and the activities they encouraged remained matters of academic concern until 1954, when the founder of the School, Léon Dostert, in cooperation with Paul Garvin and others, demonstrated with a small set of Russian sentences that a computer could actually be programmed to carry out translation, indicating in this way its capability of controlling human language. By indicating the feasibility of computers to manage texts, the demonstration brought wide attention and financial support for research in the humanities that previously seemed to require little but sabbaticals for scholars and occasional subsidies to assure subsequent publication. For somewhat over a decade the area of language study that came to be called 'computational linguistics' thereupon enjoyed broad support.

In hindsight, the basis for the manipulation is obvious and straightforward. Language is a symbolic system. The computer is a symbol-manipulating device. Yet many accomplishments that subsequently seem obvious require someone with insight, energy and daring to bring them about initially. While the basis of the control of human language by computer is straightforward, Georgetown University in its School of Languages and Linguistics achieved the first demonstration of that activity.

However great the initial enthusiasm, computer manipulation of language would not have gained academic and funding support without continuously added evidence that the expectation of success was valid. It was fortunate that the School was established with a dual aim, as indicated in its designation: 'of languages and linguistics'. Its faculty included members who controlled both language and the ability to carry out the necessary analyses, notably the organizer of this session, Michael Zarechnak. These linguists carried out the necessary and often wearisome tasks of intricate analysis that provide the

means for a mechanical device to deal with the exceedingly complex symbolic system used by human beings to communicate.

The work encountered many obstacles. Among the chief was criticism and even derision from fellow humanists. It's scarcely necessary to recall examples; but it is important to state that such responses illustrate the general ignorance of language, even among scholars who concentrate on it, such as literary specialists, and also the unawareness of its uses in society, which may be even more widespread. With a bit of thought anyone who ever bought a car might have noticed that the manual in the glove compartment is longer than works like *Paradise Regained* and many other literary classics. Moreover, manuals are not written in the style of Milton's poems, nor of his essays, nor of law briefs, nor even of newspaper articles. Attention to literary works, including their translation, might readily achieve results with no further equipment than a pen, paper and a desk. But technical materials, such as manuals for complex equipment, and texts produced by administrative organizations like governments, are infinitely longer and might well merit treatment with the help of additional equipment, both in the interests of time and completion. I might note that I didn't cover the distance from my hotel to this room by means of a 747. Much as members of society apply different tools for differing objectives, they also apply differing forms of language in different contexts for differing purposes. The promise of computerized translation to manage one of these forms--the language of technical documents--brought funding for the computerized study of language for some time after the successful Georgetown demonstration.

That study during a brief period of about a decade yielded both practical and theoretical achievements. Among practical accomplishments were translation systems that long remained in use. The Russian-to-English system developed here was installed in sites in this country and abroad, with results that met the approval of its users (see Hutchins 1986:70-78). The School also served as training center for specialists, a contribution often overlooked when evaluating centers of computational linguistics. I mention only two such specialists. Peter Toma, Georgetown's primary systems specialist, went on to develop SYSTRAN, the translation system that has been most widely applied, as by the European Community with its massive language requirements. And Muriel Vasconcellos is leader of the group that developed Spanish-English and English-Spanish systems for the Pan American Health Organization.

In spite of such achievements, the activity was discredited and the funding stopped by the report of a committee of the National Academy of Sciences, supported by the National Science Foundation. It is difficult to overstate the damage caused by this report to the national welfare and to research in the humanities. Without using time and space to detail results, I merely mention that this last year another committee of the National Academy of Sciences, its Computer Science and Technology Board, listed among six preeminent challenges a translating telephone. The committee had learned that the Japanese had appropriated 120 million dollars for a seven-year project to 'build a telephone that will translate from Japanese to English' (Waldrop 1988: 1436). The committee also understood the difficulties adequately to indicate that a necessary step on the way to the aim was achievement of machine translation. Twenty-two years after the Academy killed research and funding,

not only here but abroad as well, it suddenly resuscitated the goal, asking federal funding agencies like the National Science Foundation to provide support. In the meantime the Foundation has invested funds for investigating frog calls and other animal communication systems. We may wonder when it will assign funds to meet the revived goal.

2 Problems. The chief problem faced by linguists seeking to achieve machine translation is the widespread ignorance of language, its structure as well as applications. One can hardly read any comment on language, even by scientists successful in other areas, without finding confusion of language with its written form. And because any infant can acquire language, it is assumed to be simple. Further thought might suggest the error of such an assumption concerning a communication system embodying means to express any concept, however complex or however novel. Examples may be taken from any science, for example, biology; specialists had no difficulty devising ways for discussing the immune system, which was largely unknown before work to combat AIDS. Documentation of the immune system requires as many as a million terms. To cite one example of a well-known scientist's basic lack of understanding of language, we may recall Jerne's address on receiving a Nobel Prize. In that address he alluded to transformational grammar and proposed that research would one day uncover the biological bases of human language.

As any linguist knows, the essence of language is relationship. There is no biological basis for the word *cat*. It is associated with a value, to use Saussure's term, through its difference from other words like *sat* and *mat*. Moreover, longer segments of language such as Strawson's widely repeated sentence *The cat sat on the mat.* has its value through its difference from other sentences like *The cat sat on the mat?* or phrases like *The basset on the mat...* By identifying and mastering those values in any given language, speakers are able to communicate, to express and understand meanings. There are no biological chunks for Jerne or any future investigator to identify.

Understanding of a system based on relationships of massive complexity is elusive, as the five millennia or more of attention to language demonstrates. Much of this attention was external, such as providing lists of words with definitions, or counting letters, as in cabalistic or koranic study. In time general insights were achieved, as by Charles Sanders Peirce, identified by Sir Charles Snow as one of the two outstanding minds produced by this country in the nineteenth century.

To achieve an understanding of language Peirce posited three important relationships: those among meaningful elements of language that he called signs, those between speakers and signs, those between signs and the outside world. Such a position, as is well known, sorts out three areas of attention: the relationships among signs, which Peirce called *syntactics*; the relationships between signs and speakers, which he called *pragmatics*; the relationships between signs and the outside world, which he called *semantics*. Peirce's framework, now generally referred to as *semiotics*, permits us to distinguish and gain control of segments of human language without achieving the mastery of a normal speaker.

In this way the framework is of crucial importance for what we may provisionally call abnormal approaches to language. Among such approaches

in syntaxics is construction of artificial languages like Esperanto or Fortran. Among them in pragmatics is manipulation of a speaker's relationship to an audience, as in the PR associated with television. Among them in semantics is restriction of the sphere allowed in a limited communication system, such as expert systems. But Peirce's insights were scarcely noticed, certainly not by linguists.

A second major problem concerns support for the humanities. The field is associated with gentility. Members of humanistic professions are supposed to concern themselves with history, with literature, and other topics of intellectual but not practical benefits. Equipment, even a typewriter, might be useful for enterprising humanists but nothing that was more complex. And practical applications, like translation, scarcely achieved the dignity for discussion, let alone subsidy, except for works purified of any but academic merit. As a result, computerized attention to language remained, and remains, outside the domain recognized as proper for respectable humanists to pursue and for funding organizations in the humanities to cultivate. And linguists, notably those enjoying their self-adopted label 'mainstream', disdained the activity, except when they were able to profit from available funding.

3 Promise. To examine the promise of computational attention to language we can hardly do better than note the situation in chemistry, one of the oldest and most successful of sciences. As we all know, the field has become increasingly complex. Thirty years ago an eminent chemist of my acquaintance used to quip that his students had never used a test-tube. Some time later the department at our university set out to solidify its eminent position even further: they sought out and appointed computational chemists. As is well known today, chemists may approach no nearer to the experiments many of us carried out in our elementary chemistry course than a computer screen. And the designation of such specialists: theoretical chemists!

By contrast, in linguistics the term 'theoretical' has been appropriated by linguists who do not deal with language, but, if at all, with 'an ideal language spoken by an ideal speaker-listener in an ideal society'; preferably, they confine their attention to grammar. One may attach too much importance to designations, but I would like to propose that instead of 'theoretical' such linguists should be referred to as speculative. And attention to an 'ideal' or any other kind of mythical language should be labeled 'speculative linguistics.'

A major problem of academic or scientific attention to the humanities until recent times has been the absence of any possibility of testing one's hypotheses. The physical scientists, as Galileo demonstrated, could devise convincing tests. Even when dealing with abstruse hypotheses, like Einstein's, convincing tests were devised, as the British astronomers demonstrated to general delight by observations during an eclipse of the sun in 1918. Linguists can't look to the heavens for such demonstrations, nor can they find convincing evidence by positing fanciful structures like the Language Acquisition Device in the brain, but they can use computers. It may not be inaccurate to claim that the most important contribution of the computer is and will continue to be its provision of a device to test hypotheses formed about the means of communication developed by human beings, and thereupon to test other humanistic activities.

As above, I cannot take time to discuss the early problems. Among these were the pitifully primitive computers; it is about as pointless to try to inform a contemporary student in computational linguistics about the IBM 650 or even 709 and similar advanced computers of thirty years ago, or about punch cards for inputting data, as to portray life during the depression to a teenager today. And to mention that early computational linguists had to program in machine language puts one among the head hunters of Papua. In any event, things have changed. Anyone with any interest can learn about Rumelhart's experiments with language acquisition by computer or Sejnowski's NETtalk, a neural net 'model that can learn to pronounce English' (Roberts 1989:481). Without any inborn or implanted language acquisition device, computers have acquired elements of human language.

However admirable these experiments, we here are concerned at least in part with control of language materials for specific ends. One may ask: why translation? Why not question-answer systems? Or expert systems? Or speech signals to a robot on an assembly line? And so on. We respond to these and similar questions with the general statement that such applications make use of tricks rather than a thorough control of human language. In a sense they are clever adaptations like the computer programs that get us a seat reservation from an airline, or tell us our current bank balance. For advances in control of human language a computer has to handle it in much the breadth of its human speakers.

So-called knowledge-based systems may be even more dangerous, through specious attractiveness. They are based on the assumption that meaning exists somewhere out there, and that shrewd techniques, notably those utilizing various kinds of logic, can be used to codify and in this way control meaning. Codification is attempted through words, often with the assumption that these are the carriers of all meaning. The notion is dispelled in the first day of elementary classes in linguistics with utterances like *oh yea?* in contrast with *oh?--yea*. Devotees of knowledge-based systems might review linguistic study of the eighteenth century. To it we owe Roget's *Thesaurus*, and similar works for other languages, which are useful occasionally to writers. But the nineteenth century learned that much meaning is conveyed through phonology and morphology and syntax. Further, these meanings are fundamental to communication, as utterances like *Did she?* vs. *She did.* illustrate. In short, for advances in control of human language a computer has to handle it in much the breadth of human speakers.

To achieve that breadth, translation is the most readily realizable application, and accordingly the optimum initial goal of linguists who seek to establish a responsible base in theoretical linguistics.

I support this statement briefly: Computerized translation is concerned with technical materials, that is, with language that is most circumscribed pragmatically and at the same time most simple semantically. *Oxygen* has one meaning, in contrast with a word like *love*; and verbs like *oxidize* have a far more circumscribed meaning than do verbs like *get* or *hold*. Moreover, syntactically, technical materials are more closely regulated than is other language, even that of government documents. It is no accident that METEO, the Canadian system for translating weather reports, is one of the most rewarding that has been devised and put in use. As it and other examples

illustrate, by achieving machine translation theoretical linguistics gains control of one strand of human language. In the way of science, it will move from that mastery to broader mastery.

We might review briefly some of the aims of that mastery. In linguistics, as in other sciences, advances are most solid through a theoretical approach. Practical goals disclose problems; their solutions are most fruitful when general principles are sought, for these often solve an entire range of problems, not merely the ones detected. In view of the interruption of research, insights achieved by specialists in machine translation are forgotten or credited to other researchers. I cite only a few.

In the realm of general structure, Victor Yngve distinguished between left-branching and right-branching modifiers in languages; English restricts left-branching severely, in contrast with permitting numerous right-branching. Moreover, the two processes are favored in specific languages, as is right-branching in English--with relative clauses and object clauses placed to the right of heads-- and left-branching in other languages, like Japanese--with such clauses preceding heads, so that *the man who came to dinner* is expressed in Japanese with sequences corresponding to 'dinner to came man'. Today such study is carried out in typology, following the lead of Joseph Greenberg; specific characteristics are related to clause patterns, notably by attention to V(erb) Object) and OV languages.

In the area of pragmatics Erwin Reisler proposed distinguishing the vocabulary of what are now referred to as sublanguages or different registers. The distinction today is widely observed in sociolinguistics.

In the field of syntax we may note Zarechnak's study of the '-sja' verbs in Russian' (1971). Using twelve features, he arrives at a set of formally determined classes. It is interesting to compare his results with those of Geniušienė in her more general work 'The Typology of Reflexives' (1987), especially since both credit Xolodovic for ideas. Zarechnak's conclusions can be pursued in greater detail because of 'its suitability for computer programming', as he points out (199). Unfortunately, the cessation of funding checked the possibility of further such studies.

As a final example, Hutchins attributes the introduction of formalization into linguistics to work in machine translation (1986:59). These selected examples may illustrate the contributions we may expect for accurate understanding of language by resumption of funding.

That funding must also be adequate to support research into further areas of language and broader spheres than technical language. Some semantic studies are being carried out, but only on selected problems. The report of the Academy's Computer Science and Technology Board indicates that the way to achievement of a translating telephone requires a massive effort, including 'a machine translation system capable of dealing with all the vagaries of human language, including ambiguity, nongrammatical phrases, and incomplete phrases' (Waldrop 1988:1436). This achievement then demands computer control of the pragmatic and semantic as well as the syntactic spheres of language. The Japanese, as the Board also reports, have allotted more than a hundred million for the telephone that 'will translate from Japanese to English'; this sum is in addition to the massive amount made available over the past decade, in comparison with no funding in this country.

And as the Board recognizes, control over coded texts is a far cry from control over speech. We have a bit of work ahead of us. It is curious that the imaginative computational work on language that enjoys considerable funding is carried on not by linguists but by psychologists like the Rumelhart group and by biologists at the Salk Institute, while the primary specialists dealing with language remain without support.

And even when the translating telephone is achieved--as things are going, by the Japanese--we will be only at the beginnings of computer control over language. Translation is important, as for the European community, for scholars dealing with special problems like translation of the Bible, for forward-looking leaders in the Third World who see computerized translation as the means to bring their countries up to date in education, technology and government.

We also need to yoke the computer to identify the material in publications that is pertinent to specialists. Even in linguistics, publication is so massive that no one presumes any longer to control all of it; and chemists gave up long ago even covering the abstracts of publications in their field. We might also allude to the massive intelligence communities. For all such groups translation is only a first and partial step, far less important than control over data. Computers must be harnessed to cull out desired data, to perform data retrieval.

After data retrieval, computerized procedures will be extended to fact retrieval, that is, not simply to determine the information in texts but the knowledge controlled by human speakers. We may then be able to use the power of computers for managing the data assembled when we deal with the social and political problems of the prospective ten billion inhabitants of the earth, the data of biological complexities like those of the immune system or harmful agents like viruses, and of problems in the physical world like aerodynamic systems, as well as those resulting from experiments with supercolliders, and so on. Even with funding comparable to that provided by international research groups, we can scarcely expect such achievements until well into the next century. Yet it is scarcely surprising that the foremost countries economically today are those supporting the research and technological applications involving computational linguistic activities.

In the meantime, the control over broader types of communication, among them oral language, will yield many applications. Possibly the most welcome among these would be assistance to the handicapped. We have seen how current computer capabilities permit the noted scientist Hawking to communicate still. Control over speech, combined with advances in psychological and biological research, like the several alluded to above, could greatly assist those similarly handicapped. Other advantages of computer mastery of speech have been projected by imaginative technological specialists, including such commonplace activities as directing a vehicle. Technologists and specialists in other sciences must never forget that all of these depend on understanding language, that is, results derived from research carried out by linguists with an accurate understanding of human language.

To return to the current scene and hopes of funding, we look forward to resumption of broader activities of this university in computational studies. Thanks to Professor Zarechnak's continued work, teaching and research in this

field has never ceased here, as documented by reports like that given by him at the Kentucky conference in 1988. And as this celebration indicates, the School has maintained its energy and productiveness. Few commemorations of its important demonstration thirty-five years ago would be more significant than reestablishment of a large-scale project.

Note

Linguistics Research Center, The University of Texas, Austin, TX 78713-7247. The Center is supported under a contract with Siemens AG, Munich, and by a grant from the Division of Research Programs, Texas Higher Education Coordinating Board.

References

- Dym, Eleanor D., ed. 1985. Subject and information analysis. New York: Marcel Dekker.
- Geniušienė, Emna. 1987. The typology of reflexives. Berlin: Mouton de Gruyter.
- Hutchins, W. J. 1986. Machine translation: Past, present, future. Chichester: Horwood; New York: Wiley.
- Lehmann, Winfred P. Machine translation, in Dym 1985:110-23.
- Locke, William N. Machine translation, in Dym 1985:124-53.
- Roberts, Leslie. 1989. Are neural nets like the human brain? *Science* 242:481-82.
- Rumelhart, David E., James L. McClelland, and the PDP Research Group. 1986. Parallel distributed processing. Explorations in the microstructure of cognition. Cambridge, Mass.: MIT Press.
- Waldrop, M. Mitchell. 1988. National academy looks at computing's future. *Science* 241:1436.
- Zarechnak, Michael. 1979. The history of machine translation. In: Machine translation, by Bozena Henisz-Dostert, R. Ross Macdonald, et idem. *Trends in Linguistics* 11. The Hague: Mouton. 4-73.
- Zarechnak, Michael. 1986. The intermediary language for machine language translation. In: Computers and Translation 1:83-91.
- Zarechnak, Michael. 1988. Machine translation: Present status and future outlook in the global milieu. In: Global demands on language and the mission of the language academies, ed. John Lihani. Lexington: The University of Kentucky. 151-71.

Language teaching technology: A low-tech view

Stephen Krashen

University of Southern California

Like many readers of this paper, I am fascinated, even obsessed, by computers. I own several, I love them, and I still wistfully read the ads for the newer, improved models. My friends and I talk about hard disks, zero wait states, and 2400 baud modems. When I get up in the morning, I make some coffee, and then turn on my computer. The screen lights up, and I feel good all over.

Computers are making things possible that were impossible before--they are leading us into a new era. They are clearly one of the best things that has ever happened to the human race.

The computer clearly has something to contribute to language teaching. Computers can be used to supply comprehensible input in the students' area of interest via films, TV shows, and lectures, and researchers are working on ways of allowing students to stop the program, repeat scenes, and even interact with the input source. Computers also allow students to select appropriate input rapidly and easily (see e.g. Bush and Crotty 1989). This resource is especially useful in cases where aural comprehensible input is not plentiful.

In moving forward to take advantage of this exciting technology, however, we should not neglect another, very powerful source of comprehensible input, a source that I think has been underused in first language pedagogy and ESL, and has been nearly completely ignored in foreign language pedagogy. This source is reading.

I have attempted, in previous writings, to document 'the power of reading.' The evidence is, I think, overwhelming in first language studies and impressive in second language studies. Reading for meaning improves reading comprehension, writing style, grammar, vocabulary, and even spelling (for reviews, see Krashen 1984, 1985a, 1988). Moreover, there is good reason to suspect that reading is the only way these aspects of language are fully developed.

In this paper, I will discuss what may be the most effective kind of reading for language development. Let me introduce it with the following testimonial:

On a November day in 1957 I found myself standing in front of Miss Grosier's first grade class in Hillcrest Elementary School...trying to think of a really good word. She had us playing this game in which each kid had to offer up a word to the class, and for every classmate who couldn't spell your word, you got a point--provided, of course, that you could spell the word. Whoever got the most points received the coveted gold star.

"Bouillabaisse," said I, finally.

"You don't even know what that is," Miss Grosier scolded.

"It's fish soup."

"You can't spell that!"

"Can too."

"Come here. Write it." She demanded.

I wrote it. She looked it up, and admitted that it was, indeed, correct.

Easiest gold star I ever won. And right here, right now, I'd like to thank, albeit somewhat belatedly, whoever wrote the Donald Duck comic book in which I found the word bouillabaisse. Also, I'd like to thank my mother who read me that comic book and so many others when I was four and five....I learned to read from those sessions long before I started school. While most of my classmates were struggling with *See Spot run*, I was reading *Superman*. I knew what *indestructible* meant, could spell it, and would have cold-bloodedly used it to win another gold star if I hadn't been banned from competition after 'bouillabaisse' ...

The author of this wonderful story is Jim Shooter, former editor-in-chief of Marvel Comics; it appeared in the 1986 edition of the *Overstreet Comic Book Price Guide*.

The most effective form of reading for stimulating language development is, in my view, the kind of reading that school pretends does not exist: light reading. I suspect that light reading is the way nearly all of us learned to read; light reading may be the 'hidden curriculum' in language arts.¹

I will focus this discussion, for the most part, on comic books. Comic books are certainly not the only form of light reading around, and I am not claiming that everyone should read comics in order to develop literacy and language competency. I have chosen to concentrate on comics for several reasons. First, there have been far more studies done on the effect and content of comics than on any other form of light reading, research we shall examine below. Second, comics have been very popular: Researchers in the 1950s found that about 90% of elementary school children and 50-80% of junior high school students read comics (Slover 1959; Witty and Sizemore 1954; Blakely 1958). Third, comics are currently showing signs of increased vigor: According to a recently released survey, the number of comic book specialty shops in the United States has increased from 100 in the mid-1970s to about 4,000 in 1987, while the annual sale of comic books (not including sale of used comics) has risen from \$200 million in 1983 to \$350 million in 1987 (*Los Angeles Herald Examiner* 1987). Another sure sign of their renaissance is the fact that comic books are now on sale in general book stores in the United States, in addition to specialty shops.

Before looking at the research on comics, let us first have a look at the history of comic books in the United States. (While I have seen some excellent foreign language comic books, my data in this paper is limited to English-language comic books in the United States, because I have been unable to find research on comic books in other languages.)

1 Comics: A brief history. Comic book reading appears to be regaining acceptance. Comics enjoyed a 'Golden Age' from about 1937 to 1955, a time that saw the introduction of characters such as Superman (1938), Batman

(1939), Wonder Woman (1941), and Archie (1941). Public concern about the impact of comic books on behavior, stimulated in part by Fredric Wertham's book *Seduction of the Innocent*, published in 1954, resulted in the establishment of the Comics Code Authority, guidelines that comic book historian M. Thomas Inge referred to as 'the most severe form of censorship applied to any mass medium in the United States' (Inge 1985:12). The result was a decline: 'Writers and artists, in an attempt to "clean up their act," began to grind out boring and repetitive stories about spooks and funny animals' (Brocka 1979:30).

Despite Wertham's claims that comic book reading had a negative effect on children, research has failed to find a relationship between comic book reading and behavior (see e.g. Witty 1941, Lewin 1953; cited in Witty and Sizemore 1955, Blakely 1958. A possible exception is Hoult (1949), who reported that juvenile delinquents read more comics and more comics labeled 'harmful' and 'questionable' than did a comparison group of nondelinquents. Nearly all of Hoult's subjects, however, reported reading comic books.) Negative effects of comics, if they exist, are certainly not intrinsic to the comic format.

The recovery, the 'Silver Age' of comic books, began in 1961 with the publication of Marvel Comics' *Fantastic Four*, followed in 1962 by what may have been the most important event in comic book history in the United States: the first appearance of Spider-Man. Under Stan Lee's leadership, Marvel developed the first superheroes with problems. Spider-Man, for example, has problems that the Superman and Batman of the 1930s and 1940s never imagined--money problems, women problems, and a lack of direction and self-esteem.² Other Marvel characters are equally fascinating. Among the most popular are the X-Men, mutants with special powers who are misunderstood and mistrusted by the public.

2 Research on comic books. In my view, a good case can be made for comic books. Comic books promote language development in two ways: (1) Directly, by providing comprehensible input. The direct effects of comic books have been investigated by examining comic book texts and by researching the effects of comic book reading on language development. (2) Indirectly, by acting as a conduit to other reading.

2.1 Direct effects: Comic book texts. Contrary to popular opinion, comic book texts are well suited for language development. They contain quite a bit of language, vary in complexity (making them suitable for students at different levels of language proficiency), and contain well-formed, grammatical language.

Thorndike (1941) reported that comic book texts contain a fair amount of reading: Then currently popular comics contained about 10,000 words each. Thorndike (1941:110) commented that these figures indicate that comics are not merely picture books: 'The child who reads a comic book once a month through the school year (and this represents a very moderate dosage), gets about as much wordage of reading as he gets from even the fourth or fifth grade reader. In view of the need of the upper elementary school and junior

high school pupil for a large volume of reading and vocabulary-building experience, this source should not be neglected.'

Thorndike estimated the reading difficulty of comic book texts by means of the Lorge formula, applied to 300 word samples, and reported that the popular Superman and Batman comics were written at about the fifth or sixth grade level. More recently, Wright (1979), using the Fry formula, evaluated a wider range of comic types (Table 1). Wright's data for 'super-hero' comics (e.g. Superman, The Incredible Hulk) is quite consistent with Thorndike's 1941 findings. Wright's results also reveal a wide range of readability levels; he notes that 'humorous and funny animal comics tend to have lower readability levels' (Wright 1979:160).

Table 1. Reading level of comic books 1978 (from Wright 1979).

Comic title	Readability grade level of 3 samples			Average
The Amazing Spider Man #187 (12/78)	7.4	3.0	2.8	4.4
Archie #274 (9/78)	2.0	1.7	1.7	1.8
Batman #299 (5/78)	7.9	4.0	8.5	6.4
Bugs Bunny #201 (10/78)	2.9	1.9	1.7	2.1
Casper the Friendly Ghost #200 (10/78)	1.9	1.7	1.7	1.8
Chip and Dale #55 (11/78)	2.9	1.9	1.8	2.2
Dennis the Menace #158 (7/78)	2.8	3.0	4.7	3.5
The Incredible Hulk #74 (11/78)	5.5	9.2	1.9	5.5
Mighty Mouse # 53 (11/78)	1.9	3.3	1.9	2.4
Sad Sack #265 (11/78)	2.4	1.9	1.9	2.1
Spidey Super Stories (11/78)	2.7	1.8	1.9	2.1
Star Hunters #7 (11/78)	6.0	7.3	3.3	5.5
Star Wars #16 (10/78)	7.5	7.4	3.3	6.1
Superman #329 (11/78)	7.3	8.3	3.5	6.4
Tarzan #18 (11/78)	7.6	4.4	4.5	5.5
Tom and Jerry #311 (10/78)	1.9	2.0	1.8	1.9
Wonder Woman #245 (7/78)	5.5	5.5	3.5	4.8
Woody Woodpecker #172 (11/78)	2.4	2.4	3.0	3.1
Yogi Bear #7 (11/78)	3.2	3.5	2.4	3.0

The Fry formula is based on three samples of 100 words, taken at random. These samples can vary quite a bit. Note, for example, the variability in the three samples for *The Incredible Hulk* in Table 1 (5.5, 9.2, 1.9). Daniel Krashen has suggested to me that the 9.2 sample may have been based on the speech of Bruce Banner, the Hulk's alter ego. Banner is a research scientist, and his speech reflects his profession.

If readability scores are at all valid, Thorndike's and Wright's analyses show that comics can be at a respectable level of difficulty. By way of comparison, best-sellers in 1974 ranged in readability from grade 6 to grade

10, with a mean readability score of 7.4 (Schulze 1976, cited in Monteith 1980).

Thorndike's analysis revealed that very little of the vocabulary in comics was 'objectionable'; he found only 649 words in four comics (total = 41,000 words) that did not occur in E.L. Thorndike's *Teacher's Word Book of 20,000 Words*; of these 649 words, only 10% were classified as 'vulgar' slang (e.g. *awk, betcha, conk*). This figure underestimates the amount of slang in comics, Thorndike points out, since many standard words are used with slang meanings (e.g. *rod, joint, etc.*): 'But even when this is allowed for, one hardly gets the feeling that the language of these particular magazines was excessively slangy' (1973:112).³

To see how sophisticated comic dialogue can get, consider these examples. The first is from Marvel's *Fantastic Four*. In this scene, protagonist Reed Richards (known as 'Mr. Fantastic,' Reed Richards is a master scientist) is explaining to his wife Sue Richards ('The Invisible Woman') how the villain Psycho-Man operates:

The Psycho-Man has a vast technology at his command, darling, but he had traditionally used it to only one end: to manipulate emotions. Everything he does is designed to create conflicting, confusing emotional stimuli for his intended victims. (*The Fantastic Four*, no. 283, 1985:21)

In Marvel's *Secret Wars* no. 1, several superheroes speculate as to how they were transported to another planet:

Captain Marvel: H-how'd we get here? I mean, one minute we're checking out this giant whatchamacallit in Central Park, then 'poof,' the final frontier.

Mr. Fantastic: This much I can tell you, Captain Marvel--this device apparently caused sub-atomic particle dissociation, reducing us as we entered, to proto-matter, which it stored until it teleported us here, to pre-set coordinates in space where it reassembled us inside a self-generated life support environment!

The Incredible Hulk: That's obvious, Richards! (*Secret Wars* no. 1, p. 2)

Deborah Krashen has pointed out to me that if teachers are looking for 'high interest/low vocabulary' reading for students who need extra practice with easier texts, they can't do better than Archie! According to Wright's data, Archie is written at about the second grade level but the stories are about high school students. (By the way, after 48 years, Archie and his friends are still in high school, certainly the longest incarceration in the history of education. This is good news for students and teachers--there are lots of used Archie comics around.)

3 Comic book reading and language development. Wertham (1954) had claimed that comic book reading interfered with learning to read and with language development. The available evidence does not support these claims. In addition to Jim Shooter's testimonial, as well as other case histories we will see later, there is some educational research on the effects of comic book

reading on first language development. My interpretation of this research is that it shows that comic book reading certainly does no harm, and is probably quite helpful. We begin with two Sustained Silent Reading studies, which we will look at in some detail.

3.1 Sustained silent reading studies using comic books. Sperzel (1948) reported on a study in which three groups of 15 fifth grade students participated in sustained silent reading for six weeks, with two groups reading comics. One of the comic book groups kept a vocabulary list ('a word I liked', 'a hard word'). Sperzel reported no difference among the three groups in reading comprehension and vocabulary growth, although all three groups showed acceptable gains.

At worst, this study shows that comics are just as effective for language development as are the school activities they replaced! Note also that the duration of the study was very short. In my reviews of the effects of sustained silent reading and self-selected reading Programs (Krashen 1985a, 1988) I concluded that such programs do not show clear effects unless they last for at least seven months; programs that last seven months or longer typically result in gains in reading comprehension and vocabulary. Sperzel's study lasted only six weeks.

Perhaps the most interesting finding in Sperzel's study is how much the children enjoyed it-Sperzel (1948:111) reported that 'The period was eagerly looked forward to...as far as the rest of the world was concerned, it simply did not exist for these boys and girls.'

A second SSR study using comics was done by Arlin and Roth (1978). They randomly assigned 42 third graders into two groups. Both groups did sustained silent reading for 20 minutes daily for 10 weeks. One group chose from 130 high interest books while the other group read 'educational' comic books. It should be pointed out that this study may not have been a fair test of the benefits of comic book reading, since the comic book readers read only 'classic' type comics, which may have been less interesting than the books.

Although the duration of the study was short, both groups gained in reading comprehension--'good readers' gained similar amounts from comic and book reading, well above expected gains. While the 'poor readers' gained significantly more from book reading, they still matched expected growth when reading comic books, gaining .26 years in ten weeks, which is a little bit less than 1/4 of a school year. We can interpret these results as we interpreted Sperzel's results--if sustained silent reading time was taken out of regular English class, one could conclude that comic book reading is at least as beneficial as regular instruction. Gains of the group that made the most improvement (poor readers-regular books) were simply spectacular--one full year in 10 weeks. This is stunning confirmation of the power of free reading, and is fully consistent with other reports of the effectiveness of sustained silent reading and similar programs.⁴

4 Long-term effects of comic book reading. Long-term comic book readers, those who continue to read comic books after the early grades, show no deficit. They are at least equal to noncomic book readers in reading, language development, and overall school achievement:

- Witty (1941) compared fourth, fifth and sixth graders who claimed to be the most frequent comic book readers (the top ten percent) and the least frequent (the bottom ten percent), and found no difference in their school grades.
- Heisler (1947) found no difference in mental age, vocabulary, and reading comprehension between comic book readers and noncomic book readers in grades one through eight.
- Blakely (1958) reported no difference among non, light, moderate and heavy seventh grade comic book readers in nonverbal IQ, reading comprehension, spelling, and overall school achievement.
- Swain (1978) found that slightly more elementary school students who made good grades read comics than students who made poor grades. The reverse was found for junior high school and high school students. According to my analysis, none of these differences was statistically significant.
- Greaney (1980) reported that fifth graders who read more comic books tended to do better on a test of reading comprehension; the correlation was quite small, but was significant. Anderson, Wilson, and Fielding (1988) have recently reported similar results for fifth graders.

5 Is light reading enough? Most of the children in the studies just reviewed read both comic books and 'regular' books (see research reviewed in the next section). What about children who read only comic books? Greaney (1980) identified a group of fifth graders who were 'predominantly comic book readers,' children who did far more comic book reading than book reading (21.2% of a sample of 920 children). These children were not significantly below the average in reading comprehension, but were not as proficient as children classified as 'predominantly book readers.'

While some comics are certainly lighter than others, Greaney's results suggest that an exclusive diet of light reading will probably develop acceptable, but not advanced, levels of competence in language and literacy development.⁵

There is good evidence, however, that light reading can serve as a conduit to heavier reading. It can not only help readers develop the linguistic competence for harder reading, but can also develop an interest in books.

6 Comics as a conduit. One research study, and several case histories, show that comic book reading can lead to additional book reading.

6.1 Spider-Man at the library. Dorrell and Carroll (1981), in an article entitled 'Spider-Man at the library,' studied the effects of including comic books in a junior high school library. Comics were introduced in the library, but students were not allowed to take them out; they had to read them in the library. Dorrell and Carroll compared circulation of noncomic material and total library use during the 74 days after comics were introduced and the 57 days before they were available. The presence of comics resulted in a dramatic 82% increase in library use (traffic) and a 30% increase in circulation of non-comic books (Table 2).

Dorrell and Carroll also reported that the presence of comics in the library did not result in any negative comments from parents and that

teachers, school administrators, and library staff supported and encouraged the idea of having comic books in the library.

Table 2. Effects of including comic books in a junior high school library. (from Dorrell and Carroll 1981).

	Precomic period	Comic period
Number of students who used library ^a		
(daily average)	272.61	496.38
Circulation		
(daily average)	77.49	100.99
a. This number does not include students brought to library by teachers for class assignments.		
Precomic period: 54 days		
Comic period: 74 days.		

6.2 Some case histories. The few published case histories I have been able to find confirm my suspicion that light reading is the way many, if not most, children learn to read and develop a taste for reading.

- Kay Haugaard, a writer of children's literature and teacher at Pasadena City College in California, writes of her experiences with comic books:

As the mother of three boys who, one after the other, were notoriously unmotivated to read and had to be urged, coaxed, cajoled, threatened and drilled in order even to stay in the super slow group in reading, I wish to thank comic books for being a conduit, if not a contribution, to culture.

The first thing which my oldest boy read because he wanted to was a comic book...

Despite her initial reluctance, Haugaard bought her son comics, reasoning that 'as long as these things appealed to him where all other printed matter had failed, I let him read all he wanted. The words he learned to read here could be used in other reading material too and perhaps his skill would lure him beyond this level.'

The results were startling:

He devoured what seemed to be tons of the things...The motivation these comics provided was absolutely phenomenal and a little bit frightening. My son would snatch up a new one and, with feverish and ravenous eyes, start gobbling it wherever he was--in the car on the way home from the market, in the middle of the yard, walking down the street, at the dinner

table. All his senses seemed to shut down and he became a simple visual pipeline.

Comics did indeed lead to other reading. After a year or two, Haugaard's eldest son gave his collection away to his younger brother (who now 'pores over the comic books lovingly'), and Haugaard notes that: 'He is far more interested now in reading Jules Verne and Ray Bradbury, books on electronics and science encyclopedias.'

Haugaard's experience is consistent with the rest of the literature. Her sons' absorption in comics is identical to the reaction Sperzel's students had (see above), and the eldest son's interest in other kinds of reading agrees with the studies showing that comic book reading does not cut off book reading. (See below; it should be pointed out that the results of these studies, as well as those reviewed previously on long-term comic book reading, suggest that Haugaard's eldest son need not have given up comics in order to enjoy other books!)

- Mark Mathabane, in his autobiographical account of his youth in South Africa, mentions comic books as making an important contribution to his acquisition of English and desire to read. Mathabane had had limited exposure to English until his grandmother began to work for a friendly English-speaking family outside the impoverished ghetto where Mathabane and his family lived:

Not long after she started working for the Smiths, she began bringing home stacks of comic books: Batman and Robin, Richie Rich, Dennis the Menace, The Justice League of America, Tarzan of the Apes, Sherlock Holmes, Mysteries, Superman, The Incredible Hulk, Thor--God of Thunder, The Fantastic Four and Spider-Man.

Mathabane's reaction was similar to that of Haugaard's son:

Having never owned a comic book in my life, I tirelessly read them over and over again, the parts I could understand. Such voracious reading was like an anesthesia, numbing me to the harsh life around me. Soon comic books became the joy of my life, and everywhere I went I took one with me: to the river, to a soccer game, to the lavatory, to sleep, to the store and to school, where I would hide it under the desk, reading it furtively when the teacher was busy at the blackboard, and getting caned each time I was found out.

Mathabane credits comics with helping to bring his English to a level where he would begin to read and appreciate English books:

Midway into my eleventh year, Granny started bringing home strange-looking books and toys. The books, which she said were Mrs. Smith's son's schoolbooks, bore no resemblance whatsoever to the ones we used at my school. Their names were as strange to me as their contents: Pinocchio, Aesop's Fables and the fairy tales of the brothers

Grimm. At this point, because of reading comics, my English had improved to a level where I could read simple sentences. I found the books enthralling.

- M. Thomas Inge, a professor of Humanities, remarks that comics were clearly a conduit for him and others: 'For my generation, it was the comic book that led directly to the printed page' (1985:5). Prof. Inge has clearly not given up reading comics. His essays 'The American Comic Book' and 'EC Comic Books and Science Fiction' are informative and scholarly. This writer's experience is similar: 'I was in the low reading group in the second grade--my father encouraged comic book reading, and improvement soon followed.'⁶

7 Long-term comic reading and book reading. The popular view is that comic book readers only read comic books. The research, however, shows that in general, long-term comic book readers do as much book reading as non-comic book readers, and the results of one study (Blakely 1958) suggest they do more. While one can find 'predominantly comic book readers,' as Greaney did, such children are not typical. Researchers have used somewhat different definitions of book reading, but report similar results:

- Witty (1941) found no difference in 'general reading' between those fourth to sixth grade students who did the most comic book reading (the top 10%) and those who did the least reading (bottom 10%).
- Heisler (1947) reported no difference between the number of books owned by comic book readers and the number of books owned by noncomic book readers in elementary school.
- Bailyn (1959), in a study of fifth and sixth grade graders, found no correlation between frequency of comic book reading and book reading; comic book readers read no more books, and no fewer books, than noncomic book readers.
- Blakely (1958) found that seventh graders who read more comic books also tended to read more 'library-type books.'
- Swain (1978) surveyed children from grades 4 to 12, and reported that comic book readers 'said they had not lost interest in other books' (p. 255). Ninety-one percent of the 'good students' (grades A or B) and 79% of the poor students (D or failing) reported that they read both library books and comic books.

8 Picture reading. Before giving comics a clean bill of health, we need to consider what is perhaps the most frequently voiced concern about comic book reading. If students read comics, won't they just look at the pictures and ignore the text?

Wertham (1954), made just this accusation: The presence of pictures in comics, he maintained, actually interferes with learning to read.

If it is true that we acquire language by understanding messages, or by 'comprehensible input' (Krashen 1982, 1985b), pictures can actually help--they can provide clues that shed light on the meaning of an unfamiliar word or grammatical structure. Research confirms the potential value of pictures; Bransford and Johnson (1972), for example, have found that certain texts were

made far more comprehensible if pictures were presented along with the text (see also Omaggio 1979, Hudson 1982, Bialystok 1983).

But some comic book readers do ignore the text and look only at the pictures. Bailyn (1959) found that 27% of the fifth and sixth grade boys she observed reading comic books 'concentrated mainly on the pictures.' In Arlin and Roth's sustained silent reading study, discussed earlier, poor readers appeared to do more picture reading of comic books than good readers did.

Why are some children picture readers? At first glance, the picture reading syndrome is puzzling, since pictures do not tell the whole story in most comics, and children do not typically ignore print in their environment. I can think of several possibilities.

(1) A difficult text combined with attractive pictures. While readers can tolerate some 'noise' in texts, some incomprehensible language, too many unknown elements, will discourage any attempts at comprehension. A second grader may not even try to read the relatively complex text (and often subtle story line) of comics such as Marvel's *X-Men* or DC's *Watchmen*, but might find the pictures of great interest.

(2) Mistaken assumptions about reading. Some picture readers may be able to read substantial portions of the text but do not attempt to read. It is possible that their incorrect assumptions about reading discourage them from trying to read--because of 'reading lessons' in school, they may have the mistaken impression that in order to read, they need to know every word in the text. Such an assumption sets up a defeating sequence of events: The reader reads less, and as a result has less of a chance to acquire more language.

These are only possibilities. Frank Smith has pointed out to me that if they are true, it does not follow that picture reading can be cured by denying the reader comic books. More comic reading, not less, may be the solution; with more exposure, interest in the story might stimulate attempts at reading.

9 Conclusions on comics. The case for comics is a good one:

- Theoretically, comic books should contribute to the development of reading ability and language. The texts of comics are linguistically appropriate and pictures can help make texts comprehensible.
- Research shows that long-term comic book reading has no negative effect on language development and school achievement.
- There is suggestive evidence that comics may serve as a conduit to book reading.

It sometimes seems to me that nearly anything hi-tech and expensive that is proposed in language teaching, no matter how bizarre, is given an instant trial, and is often given instant approval without a trial. I think that we should give low-tech, inexpensive ideas at least a fair chance. Light reading requires no knowledge of DOS, no knowledge of Basic or C. And just think of how many comic books and paperback novels you can buy for the price of one 286 computer.

Appendix. The Teen Romance.

Another example of light reading that can encourage additional reading is the 'teen romance.' Parrish (1983:611) provides this characterization:

Most of the teen romance books are written to a formula. The central character is a girl, 15 to 16 years old, and the story is always told from her viewpoint. One or more boys, 17 to 18 years old are also needed. The setting is usually contemporary and familiar, such as a small town. First love is a favorite plot focus.

The joys of falling in love, the anxiety it engenders, the pain and growth of problems met, and the inevitable happy ending are all standard. However, these romances exclude sexual situations, profanity, or perversions. The conflict is usually about the heroine's feelings--insecurity, uncertainty, unpopularity, inferiority, pleasure/pain, a struggle for independence. Dialogue generally carries the action, while characterization is revealed through the romantic interaction and problems...

Like comic books, teen romances are big business; according to Sutton (1985:24), the most successful series, Francine Pascal's *Sweet Valley High*, had sold close to 10 million copies (this figure is based on 25 volumes; at the time of this writing, the series was up to number 54). Also, teen romances are read by many, if not most girls in junior high school and high school. Parrish and Atwood (1985) surveyed 250 junior and senior high school girls in three schools in the Phoenix metropolitan area, and reported that during the school year, 50% of the eighth graders said they had read from one to five teen romances, and 100% of the ninth graders had read at least five. Also, 'an astonishing 12% of the twelfth graders had read in excess of thirty novels this school year.'

While there has been little research on teen romances, the results are quite similar to those of comic book research:

- Teen romances appear to have linguistically acceptable texts, ranging from grade 4 to grade 7. *Sweet Valley Twins* is written at grade 4 reading level, *Sweet Dream Romances*, written for girls aged 10 to 15, are written at the fifth grade level, while the *Sweet Valley High* series, for age 12 and up, is written at the sixth grade level. *Caitlin*, a 'love trilogy' by Francine Pascal, ranges from grade 5 to 7 reading level. By way of comparison, recall that the mean readability level of best-sellers in 1974 was calculated to be grade 7.4. Teen romances usually contain between 40,000 and 50,000 words per volume.

- Reading teen romances does not seem to prevent other kinds of reading. Parrish and Atwood found that 'students who read the romance novels read many other kinds of literature also.'

- Teen romances seem to bring students into the library. According to Parrish and Atwood, 'Eighth and ninth graders...get their romance novels equally from friends, bookstores and school libraries. Tenth graders favor drug/grocery stores and the school library. Twelfth graders showed the most diversity: over half got their books from friends and the public library, 37% from bookstores and the school library, with little use of home and

drug/grocery stores.' Thus, despite the easy availability of teen romances, the school library still plays a significant role as a source of reading for this genre.

• There are signs that teen romances promote reading. The following, quoted by Parrish (1983), sounds very much like Haugaard's report of how comic books stimulate reading. The author is a 14-year-old girl:

I am the kind of person who hates to read, but when my mother brought home a Silhouette book for me to read, I just couldn't put it down.

Just as there has been concern about the contents of comic books, there is concern about the contents of teen romances. There has been no research on the behavioral effects of these books, but concerned teachers and parents may be interested in reading some of the more thoughtful reviews of teen romances. Sutton (1985) gives the teen romance cautious approval, suggesting that while we regard 'the lesser lights of paperback fiction as the competition,' they have some merit:

Characterization is minimal, the writing is less than graceful ("They were all being so polite and civilized the twins thought they would throw up.") and even romance is overshadowed by the soap opera suspense. But it does work: the bare-bones plots, hokey and hoary, move. The links between successive volumes are clever, and you really want to know (the way you really want to know about *Dynasty*'s Alexis Colby) what Jessica is going to pull next.

Parrish (1983) concludes that 'in spite of their drawbacks, there are many ways in which a teen romance can contribute to a balanced, sound reading program,' and suggests that teachers should:

...discuss the books with them. Let them talk with one another in groups about the books they have read. Guide their discussions by thought-provoking questions. And, most important of all, after reinforcing the reading experience, suggest a few titles which retain a strong romantic element but are a step higher on the literary value ladder.

The case for teen romances, so far, is also a good one.

Notes

1. The material that follows is a revised and expanded version of Krashen (1987).

2. Spider-Man's problems: It should be noted that Peter Parker, a.k.a. Spider-man, married Mary Jane Watson in July, 1987. At the time of this writing, Mr. and Mrs. Parker are happily married, thus solving at least one of Peter's problems.

3. Thorndike's efforts had little effect on public opinion. In the late 1940s, Zorbaugh interviewed over 3,000 adults on the subject of comic books; 40%

of those interviewed agreed with the statement 'the English used in comics is not good for children to learn.'

4. While classic comics are probably the most acceptable to parents and teachers, there is evidence that they are not all that popular with children. Wayne (1954) asked 297 seventh grade students to indicate which comic types they preferred; each student was asked to choose four from a list of 15. Classics comics ranked ninth out of the 15. When children are asked which comics they prefer, without a list to choose from, classics comics are never mentioned (for a review of these studies, see Witty and Sizemore 1954).

5. The results of several other studies suggest that language development is related to what is read. Rice (1986) reported that adults with large vocabularies and average vocabularies did about the same amount of newspaper reading and magazine reading. Those with better vocabularies, however, 'tended to read more sophisticated materials,' such as technical journals, history, literary magazines, and science magazines.

Hafner, Palmer and Tullos (1986) and Thorndike (1973) also report that better readers (ninth graders and 14-year-olds, respectively) have different tastes in reading. While these two studies do not agree in every detail, they both concluded that better readers preferred science fiction and adventure books.

6. Wood (1950) provides an additional case history in which comic strips acted as a strong motivator for a second grade reader.

References

- Anderson, R., P. Wilson, and L. Fielding. 1988. Growth in reading and how children spend their time outside of school. *Reading Research Quarterly* 23.285-303.
- Arlin, M., and G. Roth. 1978. Pupils' use of time while reading comics and books. *American Educational Research Journal* 5.201-16.
- Bailyn, L. 1959. Mass media and children: A study of exposure habits and cognitive effects. *Psychological Monographs* 73.201-16.
- Bialystok, E. 1983. Inferencing: Testing the 'hypothesis-testing' hypothesis. In: *Classroom oriented research in second language acquisition*, ed. H. Seliger and M. Long. New York: Newbury House. 104-23.
- Blakely, W.P. 1958. A study of seventh grade children's reading of comic books as related to certain other variables. *Journal of Genetic Psychology* 93. 291-301.
- Bransford, J., and M. Johnson. 1972. Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior* 11.717-26.
- Brocka, B. 1979. Comic books: In case you haven't noticed they've changed. *Media and Methods* 15.30-32.
- Bush, M., and J. Crotty. 1989. Interactive videodisc in language teaching. In: *Modern technology in foreign language education: Applications and projects*. Ed. W.F. Smith. Lincolnwood, Ill.: National Textbook Company. 75-95.
- Greaney, V. 1980. Factors related to amount and type of leisure time reading. *Reading Research Quarterly* 15. 337-57.
- Hafner, L., B. Palmer, and S. Tullos. 1986. The differential reading interests of good and poor readers in the ninth grade. *Reading Improvement* 23. 39-42.
- Haugard, K. 1973. Comic books: A conduit to culture? *The Reading Teacher* 27. 54-55.
- Heisler, F. 1947. A comparison of comic book and non-comic book readers of the elementary school. *Journal of Educational Research* 40.458-64.
- Hoult, T. 1949. Comic books and juvenile delinquency. *Sociology and Social Research* 33.279-84.
- Hudson, T. 1982. The effects of induced schemata on the 'short circuit' in L2 reading: Non-decoding factors in L2 reading performance. *Language Learning* 32.23-43.

- Inge, M.T. 1985. *The American comic book*. Columbus, Ohio: Ohio State University Library.
- Krashen, S. 1982. *Principles and practice in second language acquisition*. New York: Prentice-Hall.
- Krashen, S. 1984. *Writing: Research, theory and applications*. New York: Prentice Hall.
- Krashen, S. 1985a. *Inquiries and insights*. Hayward, Calif.: Alemany Press.
- Krashen S. 1985b. *The input hypothesis: Issues and implications*. New York: Longman.
- Krashen, S. 1987. *Comic book reading and language development*. Victoria, British Columbia: Abel Press.
- Krashen, S. 1988. Do we learn to read by reading? The relationship between free reading and reading ability. In: *Linguistics in context: Connecting observation and understanding*. Ed. D. Tannen. Norwood, N.J.: Ablex. 269-98.
- Los Angeles Herald Examiner. 1987. Comic book industry matures as readers come of age. October 4, D4, D6.
- Mathabane, M. 1986. *Kaffir boy*. New York: Plume.
- Monteith, M. 1980. How well does the average American read? Some facts, figures and opinions. *Journal of Reading* 20.460-64.
- Omaggio, A. 1979. Pictures and second language comprehension: Do they help? *Foreign Language Annals* 12. 107-16.
- Parrish, B. 1983. Put a little romantic fiction into your reading program. *Journal of Reading* 26. 610-15.
- Parrish, B., and K. Atwood. 1985. Enticing readers: The teen romance craze. *The California Reader* 18. 22-27.
- Pascal, F. *Sweet Valley High*. New York: Bantam.
- Pascal, F. *Caitlin*. New York: Bantam.
- Rice, G.E. 1986. The everyday activities of adults: Implications for prose recall--part I. *Educational Gerontology* 12. 173-86.
- Shooter, J. 1986. *Marvel and me*. In: *The comic book price guide*, 16th edition. Ed. R. Overstreet. New York: Harmony Books. A85-96.
- Slover, V. 1959. Comic books vs. story books. *Elementary English* 36. 319-22.
- Sperzel, E. 1948. The effect of comic books on vocabulary growth and reading comprehension. *Elementary English* 25. 109-13.
- Sutton, R. 1985. Librarians and the paperback romance. *School Library Journal* 32. 25-29.
- Suzanne, J. *Sweet Valley twins*. New York: Bantam-Skylark.
- Swain, E. 1978. Using comic books to teach reading and language arts. *Journal of Reading* 19.198-200.
- Thorndike, R.L. 1973. *Reading comprehension education in fifteen countries*. New York: John Wiley.
- Wayne, R. 1954. Survey of interest in comic books. *School Activities* 25.244.
- Wertham, F. 1954. *Seduction of the innocent*. New York: Rinehart.
- Witty, P. 1941. Reading the comics: A comparative study. *Journal of Experimental Education* 10.105-09.
- Witty, P., and R. Sizemore. 1954. Reading the comics: A summary of studies and an evaluation, I. *Elementary English* 31.501-06.
- Witty, P., and R. Sizemore. 1955. Reading the comics: A summary of studies and an evaluation, III. *Elementary English* 32.501-06.
- Wood, A. 1950. Jimmy reads the comics. *The Elementary School Journal* 51. 66-67.
- Zorbaugh, H. 1949. What adults think of comics as reading for children. *Journal of Educational Sociology* 23.225-35.

Appendix: Three historical notes

G. Richard Tucker

Center for Applied Linguistics

Michael Zarechnak

Georgetown University

Given the commemorative character of this Bicentennial year and the fact that the *Georgetown University Round Table on Languages and Linguistics 1989* marks anniversaries of the Center for Applied Linguistics, the Georgetown University-IBM Machine Translation Project, the Georgetown University School of Languages and Linguistics, and the Round Table itself, these three historical notes may be of interest.

The Center for Applied Linguistics. The Center for Applied Linguistics (CAL) began operation on Monday, February 16, 1959. Participants at a conference held in 1958 at the University of Michigan had concluded that there was a need for an organization which could address certain language problems that, in the late 1950s, were becoming increasingly urgent: for example, the problem of how to meet the rapidly increasing demand for the teaching of English abroad; the problem of training enough Americans in the major languages of Asia to overcome the serious Asian-American language barrier, and the problems faced by many speakers of so-called nonstandard versions of English in the United States--to name a few. Thus, to address these problems and challenges, CAL was begun initially with funding from the Ford Foundation, and for administrative purposes, lodged under the umbrella of the Modern Language Association.

Since its inception, CAL has had four Directors: Charles Ferguson of Stanford University (who is with us here at *GURT*), John Lotz, Rudolph Troike, and currently Richard Tucker.

Many of the activities and emphases begun by Charles Ferguson continue to this day--namely, a commitment to the systematic collection and dissemination of research results and other information intended to improve educational practice (in this regard, for example, CAL has operated the ERIC Clearinghouse on Languages and Linguistics for more than 20 years); special attention to problems of the teaching and study of the less commonly taught languages (CAL continues to maintain probably the most extensive database of this kind in the world, with information about the teaching of more than 1,000 of the world's languages); and attention to the description of what we now refer to as nonmainstream varieties of language (e.g. Black Vernacular English, Puerto Rican English, Native American English, Vietnamese English, etc.). It was also during these years that CAL staff collaborated with other international universities and organizations to conduct a series of national language surveys such as the five-country survey of language use and language

teaching in East Africa. During this time, CAL also participated with four other organizations in the planning meetings which led to the formation of TESOL.

Under the leadership of John Lotz, attention was turned to expanding contacts with linguists and language educators in Eastern Europe and a series of collaborative projects were begun which resulted in the publication of contrastive analyses resulting from work with colleagues in Hungary, Poland, Roumania and Yugoslavia.

Upon assuming the directorship in 1973, Rudy Troike turned CAL's attention to two newly emerging critical sets of issues--namely, meeting the educational needs of language-minority, limited-English proficient, youngsters here in the United States; and near the end of his tenure, with the fall of Saigon, marshalling CAL's resources to help meet the needs of American educators and social service providers who would be called upon to assist the waves of Indochinese refugees soon to be arriving. Toward this first goal, CAL staff worked for many months to help the San Francisco Unified School District develop a compliance plan in response to the *Lau vs. Nichols* Supreme Court decision. With respect to the second goal, CAL initially through its National Indochinese Clearinghouse and Technical Assistance Center and more recently through its Refugee Service Center has played and continues to play a major role in providing information and meeting the needs of such newcomers.

During the past several years under Dick Tucker, CAL has become increasingly concerned with issues related to workplace communication (e.g. addressing the literacy, numeracy, problem-solving and decision-making skills of the workforce of tomorrow); has worked to facilitate full and effective participation by language-minority individuals in science and mathematics education; and has worked together with JNCL and other organizations to develop what Dick Tucker referred to in a presentation at *GURT 1985* as a language competent American society.

--G. Richard Tucker

GU-IBM Machine Translation Project. It was 35 years ago, on January 7, 1954, when the results of the GU-IBM experiment were announced in New York. The announcement was given wide publicity in the press and it elicited a wealth of comments from all quarters. Some people were astonished that machine translation was possible at all. Some people were disappointed that the quality of the translation was far from polished English prose. But those who understood the difficulties of machine translation realized that a definite and decisive first step had been taken along a road that still wound far ahead before the final goal of acceptable machine translation could be reached.

Early in 1956 the Soviet Union announced a successful translation of English into Russian. They acknowledged the relationship between their undertaking and the GU-IBM experiment. In June 1956, GU received a substantial grant from the National Science Foundation to undertake intensive research for the translation of Russian scientific texts into English.

The General Analysis Techniques (GAT) group headed by Michael Zarechnak proceeded in the research one sentence at a time. In this approach each sentence is analyzed into translation units whose presence, absence, and positional relationships to each other are all important. The analysis is carried

out at every possible level that will elicit useful information. Word formation (morphology) is the first level and includes word collocation (idiom). Word grouping (syntagmatic processes) is the second level. This includes the agreement of adjectives with nouns, the government of nouns by verbs, or other form-classes, and the modification of adjectives, verbs and other adverbs by adverbs. The organization of word groups into sentences (major syntax) is the third level. This is specifically the relationship of subject to predicate. The possibility that there are other levels is not precluded.

In the 1960s machine translation (MT) was primarily the concern of the university linguists and computer engineers. The dramatic miniaturization and the increasing sophistication of high technology have changed the picture in the 1980s. The research procedures characteristic of GAT were further developed by other MT systems in both pragmatic terms such as PAHO and theoretical such as Eurotra. Soviet systems include AMPAR, NERPA, FRAP. Japanese MT projects too are now part of larger research efforts.

All these various types of MT systems have their own various places and functions. The MT field itself is a mixture of practice and research. Experimental systems test new ideas both on large-scale computers and microcomputers. The AI methods, new parsers, logico-semantic representations are entering the field via the well-tested methods developed for MT in the past.

After a decade of disenchantment with the idea of useable machine translation, resulting to a large degree from the ALPAC Report in 1966, an upsurge in interest in machine translation has been noted recently in many countries, including the United States. The reason for the rebirth of interest in machine translation may be partly due to the progress made in recent years in the utilization of computer technology in various areas of information transfer and word processing which were unknown just a few years ago. In view of the growing volume of foreign documentation on the subject, untranslated but potentially valuable, a return to machine translation may be not only feasible but necessary. A reassessment of the whole machine translation issue is long overdue.

--Michael Zarechnak

Georgetown University School of Languages and Linguistics. Born of the desire to provide young men and women with the linguistic and cultural skills necessary for communicating in the modern world, the Institute of Languages and Linguistics was founded in 1949 by the Rev. Edmund A. Walsh, S.J., as part of the School of Foreign Service. Its first Director, Professor Leon Dostert, who was born in France in 1904 and, as a boy during World War I, served as interpreter for the American soldiers, during World War II was the official interpreter to General Giraud and to General Eisenhower, and, after World War II, was Chief of the Language Division of the Nuremberg War Crimes Tribunal. In this capacity, he brought to the world's attention the system which is now known as simultaneous interpretation.

The opening of the Institute drew international attention for two of its major innovations: the multilingual room and the language laboratory. The former permitted simultaneous interpretation in five languages; the latter introduced recent technological and electronic advances such as tape-recorders

and other audiovisual aids in the field of language teaching. Later experiments included drill exercises for the laboratory and for language testing. With the aid of IBM, Professor Dostert went on to conduct research on practical machine translation.

In 1950 Professor Dostert launched the annual Georgetown University Round Table on Languages and Linguistics, and subsequently the series of monographs which report on the Round Table each year. Held during Easter break, the first meeting drew so few people, the story goes, that they fit with ease around one table. The number of participants jumped to 144 in 1951 and has increased steadily since then. With between 350 and 500 participants attending the Round Table, scholars, teachers, and students now come from such places as Australia, Great Britain, Canada, France, Finland, West Germany, Brazil, Mexico, Japan, and the USSR, and from all corners of the United States.

The School of Languages and Linguistics, as it is known today, went on to develop Applied Linguistics, which grew out of the earlier conception of the Institute of reinforcing the teaching of languages with research in linguistics. Linkages with government agencies, the armed services, and the foreign service have influenced, from the outset, the teaching of languages at Georgetown University, which is characterized by a strong sense of communicative competence.

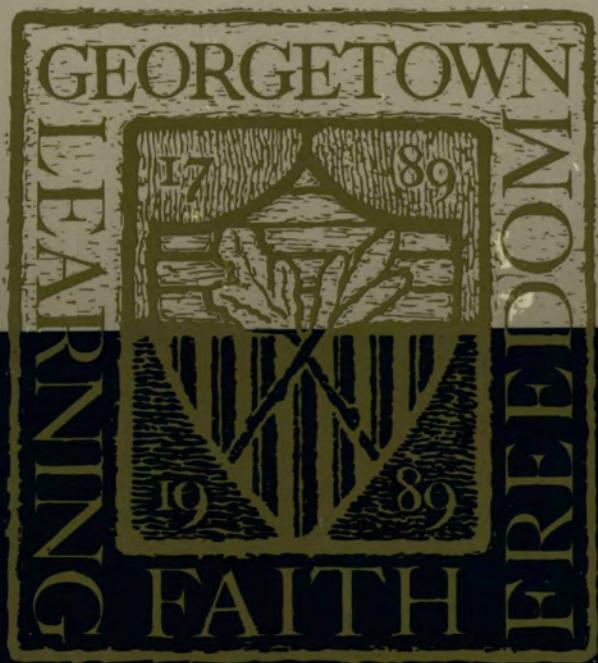
Professor Robert Lado, recognized internationally for his major contributions to the field of language testing, succeeded Professor Dostert in 1960, and later became the first dean of the new School of Languages and Linguistics. Professor Lado served as Dean until 1973.

Since 1966, when Professor James E. Alatis became Associate Dean, after a career in federal service at the Department of State and the U.S. Office of Education, and later after he assumed the deanship of the School in 1973, he has maintained the same traditions and promoted intercultural and interdisciplinary studies. Research on the interaction of linguistics, language teaching, and technology is still a major focus of the School.

The School of Languages and Linguistics has nine departments: Linguistics, Arabic, Chinese and Japanese, French, German, Italian, Portuguese, Russian, and Spanish, and three divisions: the American Language Institute, English as a Foreign Language, and Interpretation and Translation. Since 1973 the unified Department of Linguistics distinguishes four graduate programs: Applied, Computational, Theoretical, and Sociolinguistics. Its graduate program is the largest in the University.

The School is also the administrative home for the Georgetown University Press which each year publishes the proceedings of the GURT, as well as many other works on language, literature, and linguistics, along with other scholarly volumes in such fields as ethics, philosophy, and theology.

From modest beginnings the Round Table has flourished, examining issues in sociolinguistics, bilingual education, second language teaching, world Englishes, and language in public life. The emphasis of the 1989 Round Table is language teaching, testing, and technology and, in many ways, reflects the dreams and aspirations of the pioneers of Languages and Linguistics at Georgetown.



Bicentennial

0-87840-124-5