



UNIVERSITY OF  
ILLINOIS LIBRARY  
AT URBANA-CHAMPAIGN



**CENTRAL CIRCULATION BOOKSTACKS**

The person charging this material is responsible for its return to the library from which it was borrowed on or before the **Latest Date** stamped below.

Theft, mutilation, and underlining of books are reasons for disciplinary action and may result in dismissal from the University.

TO RENEW CALL TELEPHONE CENTER, 333-8400

UNIVERSITY OF ILLINOIS LIBRARY AT URBANA-CHAMPAIGN

JAN 0 4 1994

AUG 2 3 1994

JAN 0 5 1998

NOV 2 5 1997

OCT 0 7 2005

NOV 3 0 ENT'D

When renewing by phone, write new due date below previous due date.

L162

Digitized by the Internet Archive  
in 2011 with funding from  
University of Illinois Urbana-Champaign

<http://www.archive.org/details/teslstudies34univ>





# *TESL* *STUDIES 1980*



THE LIBRARY OF THE

MAY 11 1983

UNIVERSITY OF ILLINOIS  
AT URBANA

3

*PUBLICATION OF THE DIVISION OF ENGLISH AS A SECOND LANGUAGE  
UNIVERSITY OF ILLINOIS  
URBANA, ILLINOIS*

# *TESL STUDIES*

*EDITORS: YAMUNA KACHRU, J RONAYNE COWAN*

*EDITORIAL BOARD: Katherine O. Aston, Lyle F. Bachman,  
Lawrence F. Bouton, H. Douglas Brown, Lonna J. Dickerson,  
Wayne B. Dickerson and Rebecca G. Dixon*

*TESL STUDIES is intended as a forum for the presentation, in pre-publication form, of the research done by the faculty and students of the Division of English as a Second Language, University of Illinois, Urbana-Champaign. TESL STUDIES will also publish research of other University of Illinois faculty and students if it is of relevance to our field. All interested in obtaining copies of TESL STUDIES should write to the following enclosing payment (Checks should be made payable to the University of Illinois): Mrs. Norma H. Robinson, Division of English as a Second Language, 3070 Foreign Languages Building, Urbana, Ill. 61801.*

*PUBLICATION OF THE  
DIVISION OF ENGLISH AS A SECOND LANGUAGE  
UNIVERSITY OF ILLINOIS  
URBANA, ILLINOIS*

*Price per copy: \$4.00*



420.1  
E157  
V. 3-4

# ***TESL STUDIES***

## **EDITORS**

**YAMUNA KACHRU  
J RONAYNE COWAN**

## **EDITORIAL BOARD**

*Katherine O. Aston; Lyle F. Bachman  
Lawrence F. Bouton; H. Douglas Brown  
Lonna J. Dickerson; Wayne B. Dickerson  
Rebecca G. Dixon*

**VOLUME 3  
1980**

***DIVISION OF ENGLISH AS A SECOND LANGUAGE  
UNIVERSITY OF ILLINOIS  
URBANA-CHAMPAIGN***



## CONTENTS

Lyle F. Bachman	The construct validation of oral proficiency tests	1
H. Douglas Brown	The role of teacher feedback in preventing the fossilized errors of second language learners	21
Gary A. Cziko	Psychometric and edumetric approaches to language testing: implications and applications	33
Wayne B. Dickerson	A pedagogical interpretation of generative phonology I. Theoretical foundations	63
Cynthia A. Gockley	Toward developing an optimal test of writing proficiency	99
Braj B. Kachru	Models for new Englishes	117
Yamuna Kachru	'Transfer' in 'Overgeneralization': Contrastive linguistics revisited	151
Susan Osuch-Hatziavramidis Lonna J. Dickerson	Phonological and morphological conditioning of {Z} in the speech of Japanese ESL students	163
Richard Bruce Rickard	First language perceptual strategies and reading in a second language	183

## REVIEWS

Barbara Matthies	H. Douglas Brown: Principles of language learning and teaching	199
Mary Reichardt	Jerrilou Johnson: Living language: USA culture capsules for ESL students	203



## THE CONSTRUCT VALIDATION OF ORAL PROFICIENCY TESTS

Lyle F. Bachman

Adrian S. Palmer

The validation of oral language proficiency tests has long been problematical. The most common approaches to this problem have involved concurrent validation procedures for relating "indirect" measures to so-called "direct" measures. The most frequently used criterion in such studies has been the Foreign Service Institute (FSI) oral interview. In validating the FSI interview itself, however, criterion-referenced procedures are not possible, since an adequate operationally defined criterion for comparison is not likely to be found, and construct validation procedures are necessary.

In this study, a multitrait-multimethod matrix, comprising six measures, representing combinations of two traits (speaking and reading) and three methods (interview, translation, self-rating), was used to examine the construct validity of the FSI oral interview. The six measures were administered individually to 74 native Mandarin Chinese speakers of English. Data were analyzed using the Campbell-Fiske criteria for convergent and discriminant validity, and using confirmatory factor analysis. The results indicate both convergent and discriminant validity for the FSI interview. Furthermore, of the three methods used to measure oral proficiency, the FSI interview had the largest trait component and the smallest method component. More importantly, from a theoretical standpoint, the results provide strong evidence for the distinctness of the two traits, speaking and reading, and thus support the divisible trait hypothesis of language proficiency.<sup>1</sup>

### INTRODUCTION

One of the most persistent areas of difficulty in language testing continues to be the measurement of oral proficiency. Both the validity of tests which claim to measure this construct and the distinctness of the construct itself have recently been the objects of considerable research. (Clark, 1975, 1978, 1979; Palmer and Groot, 1980; Oller and Perkins, 1980) Much of the research into the validity of oral tests has employed concurrent validation procedures for relating "indirect" tests to "direct" tests, which have greater "appearance" of validity. The most frequently used "direct" criterion for such studies has been

the Foreign Service (FSI) oral interview, or some variation thereof. The FSI interview itself, however, "has far outrun its verified technical validity as a measure" (Stevenson 1980), and while many researchers believe that the FSI interview is a valid measure of "real life" proficiency, such rationalization is hardly a demonstration of validity. As the FSI interview technique is being used increasingly outside the controlled environment of FSI and in a wider variety of educational settings, the need for evidence demonstrating its validity has become more critical.

The necessity for construct validation, particularly in the absence of a fully valid criterion, has been extensively discussed (Cronbach, 1971; Cronbach and Meehl, 1955). Construct validity refers to the extent to which a test, or set of tests yield scores which are related in ways predicted by a particular theory of constructs. To investigate construct validity, one develops a construct (a theory), which becomes a provisional explanation of test results until the theory is falsified by the results of testing hypotheses derived from it. Clearly, since the FSI interview is the criterion against which the majority of other oral proficiency tests have been validated, an adequate operationally defined criterion for its validation is unlikely to be found, and hence construct validation is indicated.

While construct validation can demonstrate the validity of a given test, the results of this procedure may also provide information on the nature of the constructs, or traits, themselves. Specifically, the results of construct validation studies of language tests may provide the best means for investigating the extent to which the traditional language skills of listening, speaking, reading and writing, for example, are unitary or divisible traits. The results of this study, then, pertain not only to the construct validity of the FSI oral interview, but also to the validity of the traits, speaking and reading.

#### METHOD

The construct validation paradigm used for this study is the classic multitrait-multimethod matrix, first described by Campbell and Fiske in

1959. This paradigm has two distinct advantages over other validation procedures. First, it recognizes two types of validity, convergent and discriminant, and enables the investigator to examine both. Convergent validity is the extent to which different measures of the same construct tend to agree in their results, and is the type of validity sought in criterion-referenced validation. Discriminant validity, on the other hand, is the extent to which measures of different traits, but which use the same method of measurement, tend to disagree in their results. Convergent validity is indicated by high correlations among different measures of the same trait, while discriminant validity is indicated by low correlations among similar methods for measuring different traits. The second advantage of the multitrait-multimethod paradigm is that it allows the researcher to distinguish the effect of measurement method from the effect of the trait being measured. The influence of method on test results has been demonstrated in a number of studies (Clifford, 1973, 1980; Corrigan and Upshur, 1978; Brüttsch, 1979), and construct validity studies which ignore the effect of method are not likely to yield interpretable results.

In addition to the correlational analysis of the Campbell-Fiske model, confirmatory factor analysis (Jöreskog, 1969) was used to examine the factor structure of the measures. Confirmatory factor analysis makes possible tests of statistical significance for comparing relationships predicted by a given causal model with those observed in empirical data. Specifically, this analysis was used to compare the explanatory power of two models--one trait and two traits--by applying a statistical test for the goodness of fit of each model to the data.

#### Instrumentation

In order to test specific hypotheses regarding convergence, discrimination and the effect of method, the multitrait-multimethod paradigm requires that at least 2 traits (multitrait) be measured by at least 2 methods (multimethod). The two traits identified for this study were oral proficiency in English and reading comprehension in English. Three methods were selected: the "interview" method, the "translation" method

and the "self-rating" method. These are described briefly in Figure 1, and are discussed below.

TRAITS	METHODS		
	"Interview"	"Translation"	"Self-Rating"
Speaking	FSI Oral Interview	The subject translates replies to questions or directives written in his native language into spoken English and records his translation.	The subject rates his own speaking ability on a scale similar to that used by FSI examiners.
Reading	An interview in the subject's native language. The subject reads passages and the examiner asks the subject questions about the meaning of the passages. Both the questions and the responses in the subject's native language.	The FSI reading test, administered <u>not</u> as an interview, but as follows: the subject is given a set of graded passages in English to translate line by line into his native language.	The subject rates his own reading ability on a scale similar to that used by FSI examiners.

Figure 1. TRAIT AND METHOD MATRIX

### The Interview Test of Speaking<sup>2</sup>

The FSI interview was selected because it is widely used, and is the subject of considerable interest and controversy in the field of language testing. In addition, it is described in detail in the literature, and is operationally compatible with our provisional definition of the speaking trait.

### The Interview Test of Reading

An interview format was developed for testing reading comprehension



in which the subject was given a short passage in English to be read silently. The subject was then asked a number of questions about the passage, both these questions and the subject's answers were in the subject's native language. The questions were of the following five types, none of which required the subject to translate directly from the English passage into his native language.

- 1) Questions to be answered by pointing to information in the passage
- 2) Yes-no questions
- 3) Questions asking for a summary of part or all of the passage
- 4) Questions requiring comprehension of particular words or phrases
- 5) Questions requiring comprehension of the organization of the passage

The passages were selected according to the criteria set out in the Foreign Service Institute Testing Kit (Adams and Frith, 1979). These criteria specify the types of sources for reading passages at the five FSI levels and constitute a procedure for grading passages as to difficulty. At one extreme, the readings consisted of individual signs (of one to three words and/or numeral groups), such as street signs and signs one would encounter in a building, and a short passage of the type found in beginning language textbooks. At the other extreme, the readings consisted of a passage by I. A. Richards, a humorous piece by Phyllis Diller, and a handwritten text.

#### The Translation Test of Speaking

The translation test of speaking was adapted from the recorded oral production test (ROPE) developed by Clifford and Lowe (Lowe and Clifford 1980). The ROPE consists of a set of recorded questions or directives at FSI level 1-4. Question types at each level follow the guidelines set out in the Handbook on question types and their use in LLC oral proficiency tests (Lowe, 1976). In the original ROPE, the subject listens to the question and is given time to respond. The tape is then rated as per FSI guidelines.

For this study, the ROPE test was converted into a recorded oral

translation test (ROTE) by supplying the subject with a written Chinese translation of an answer to the recorded question/directive. This translation incorporated grammatical structures, markers of textual cohesion, and lexical items consistent with the descriptions of competence at the FSI level for which the eliciting question/directive was prepared. Thus, the ROTE test as used consisted of the following steps:

- 1) The subject listened to a tape recording in which he/she heard a question or directive, followed by an appropriate response, and at the same time read the question/directive and response in Chinese.
- 2) After a brief pause the response was repeated orally.
- 3) The subject was then given a period of time to prepare an appropriate translation into English.
- 4) Upon signal, the subject then translated the response into English, recording it on tape.
- 5) This procedure was repeated for all of the questions at each of the four levels.

#### The Translation Test of Reading

The procedure used by the FSI for testing reading was selected. Though called an interview, the FSI reading test is actually a translation test. In the FSI version of the test, the subject sits down with two examiners, who provide a short reading passage in the language being tested, and the subject translates the passage orally into his/her native language. This oral translation is rated by the examiners, and depending on the adequacy of the translation, the subject is then given a higher level or lower level passage, or a passage from the same level.

The test used in this study involved no face-to-face interaction between examiner and subject. The examiner merely handed the subject a passage to translate, determined whether the subject should be given a second passage from the same level or from a higher or lower level, and supervised the recording of the subject's translations. Since the FSI does not test proficiency in reading English as a second language by means

of their translation test, passages had to be selected for this purpose. Passages used were generally different sections from the same sources used in the interview test of reading; the few exceptions were of very similar type and difficulty.

### The Self-Rating Tests

The self-rating tests of speaking and reading were questionnaires adapted from the FSI Testing Kit written in Chinese. Each contained two different types of questions. One type probed the subject's perception of his/her functional control of spoken or written English. In these questions, the subject was asked what he/she could do with the language-- what language use situations he/she could cope with. The situations were drawn from the functional portions of the FSI guidelines and were grouped according to FSI level. The second type of question probed the subject's perception of his/her control of linguistic forms (range and accuracy). These levels of control were also drawn from the FSI descriptions of the five levels of competence. Subjects' responses to each question were either "yes" or "no."

### Background Questionnaire

Though not part of the validation study per se, a background questionnaire was administered to obtain demographic information about the subjects. It was adapted from a questionnaire which had been used previously in studies conducted at Brigham Young University. Questions were modified to make them more relevant to the subjects in the study.

All tests were informally pretested on a small group of native Mandarin speakers who were excluded from the study itself. Test procedures and items were modified as required.

### Sample

In order to facilitate administration of tests involving translation, it was decided early in the study to sample subjects from a homogeneous native language background. The particular group identified was native

Mandarin Chinese speaking students at the University of Illinois, Urbana-Champaign. Subjects were contacted at random, using a list of Chinese students obtained from the International Student Office, University of Illinois. In order to increase the variability of the sample, student spouses were also asked to participate. 85 subjects were scheduled for testing and sent background information questionnaires. Of the 85 scheduled for testing, four did not show up, two were eliminated because their control of Mandarin was not sufficient for them to complete the translation tests, and four were eliminated because they were missing one of the tests. Subjects were paid for their participation.

The subjects were 75 native Mandarin Chinese speakers from Taipei who were living in Illinois. 61 were university students (57 graduate, 4 undergraduate) majoring in 39 different fields, 13 were spouses of students, and one was enrolled in an intensive English institute. There were 39 females and 36 males, ranging in age from 19 to 35 years, with a median age of 26 years. 25 had been living in the U.S. for less than one year, while 50 had been living in the U.S. for one year or more. All had studied English for at least one year in Taiwan, and 61 had studied English for more than one year here. 30 indicated that they knew languages other than English (French: 5, German: 10, Japanese: 13, Spanish: 2 and Malay: 1). Of these, only one indicated a better knowledge of speaking and reading this language (Malay) than in English.

## Procedures

### Test Administration

Each subject took all tests in sequence, in a two-hour period. Although there was some variation in order, due to administrative constraints, the sequence generally followed was 1) background questionnaire, 2) self-rating, speaking, 3) self-rating, reading, 4) recorded oral translation examination (ROTE), 5) reading, interview, 6) reading, translation, and 7) oral interview. All tests were administered individually by project staff, and all but the self-appraisals were recorded for later scoring.

### Scoring

Each of the two interviewers administering the oral interview assigned an independent FSI rating (0 - 5 scale) to each subject immediately upon completion of the interview, after which a joint "conference" rating was assigned for use in the analysis. For the reading interview and reading translation test, each interviewer assigned an independent FSI rating to each subject. These two interviewers then rated each other's tapes, providing two sets of ratings for each measure. An average rating for each subject was computed for use in the analysis. The tape recordings of the ROTE were rated independently by two raters, and average ratings computed for use in the analysis. Scores for the two self-appraisals were the total number of questions answered "yes" by each subject on each measure.

### Analyses

Distributions, correlations and reliabilities were computed using SPSS Version 8, on the CYBER system at Illinois. Maximum likelihood confirmatory factor analyses were computed using LISREL-4, also on the CYBER system at Illinois.

## RESULTS

### Reliabilities

Because of the varied nature of the measures used, no single reliability estimate was appropriate for all. Estimates based on variance components of scores were, however, computed for all tests. For the ratings (oral interview, reading interview, reading translation and ROTE), the intraclass correlation was used, and for the self-ratings, Guttman's lambda 6, a lower bounds estimate, was used. In addition to these estimates, both inter- and intra-rater reliabilities were estimated to determine the stability of the ratings across raters and across time. The obtained reliability estimates are given in Table 1.

TABLE 1  
Reliability estimates for trait-method units

	Oral Interview	Reading Interview	ROTE	Reading Translation	Speaking Self-rating	Reading Self-rating
Inter- rater (N = 75)	.882	.975	.787	.944	NA	NA
Intra- rater (N = 30)	--	.984	--	.997	NA	NA
Intra- class (N = 75)	.878	.974	.860	.944	NA	NA
Alpha (N = 75)	NA	NA	NA	NA	.908	.851
Lambda 6 (N = 75)	NA	NA	NA	NA	.959	.894

NA = Not appropriate

-- = Not computed

### Correlations

The intercorrelations of scores on tests used in this study are presented in Table 2. Of the six tests administered, four (the interview tests of speaking and reading and the translation tests of speaking and reading) were rated by two different examiners. For the purpose of our analysis, we have considered each examiner's ratings as a separate method (or a separate test). Thus, Int-1 in Table 2 stands for the interview as rated by interviewer number 1. Int-2 stands for the interview as rated by interviewer number 2, and so on. Considering the data in this way provided us with a 10 x 10 matrix of intercorrelations. In considering the hypotheses of convergent and discriminant validity, three types of correlations are of particular interest: 1) correlations between different

measures of the same trait (monotrait), or validity indices, 2) correlations between measures of different traits, but using the same method (monomethod), and 3) correlations between measures which have neither trait nor method in common (heterotrait-heteromethod). In Table 2, the validity indices are enclosed in the triangles in the upper left and lower right hand corners of the matrix. The monomethod correlations are indicated by the diagonal dashed lines (method diagonal). All the other correlations are heterotrait-heteromethod correlations.

TABLE 2

MTM Correlation Matrix

All correlations sig. at  $p < .01$ ,  $df = 74$

		Speaking (A)					Reading (B)				
		Int-1 (1)	Int-2 (2)	Trans-1 (3)	Trans-2 (4)	Self (5)	Int-1 (1)	Int-2 (2)	Trans-1 (3)	Trans-2 (4)	Self (5)
A	1	1.00									
	2	.63	1.00								
	3	.77	.72	1.00							
	4	.76	.72	.85	1.00						
	5	.51	.56	.46	.53	1.00					
B	1	.54	.45	.62	.65	.58	1.00				
	2	.56	.46	.64	.67	.60	.97	1.00			
	3	.58	.61	.64	.68	.46	.65	.65	1.00		
	4	.52	.55	.62	.69	.49	.65	.65	.94	1.00	
	5	.44	.45	.47	.51	.68	.68	.68	.54	.54	1.00



### Convergent Validity

The first hypothesis examined concerns convergent validity, and states that correlations between scores on tests of the same trait which employ different methods (validity indices) should be significant and positive. These validity indices are all significant and positive, thus providing evidence of convergent validity for both the speaking and the reading tests.

### Discriminant Validity

The first discriminant validity hypothesis is that the validity indices should be higher than correlations between measures having neither method nor trait in common. For example, compare the validity indices in the left column of the upper left triangle (.88, .77, .76, and .51) with correlations between measures which share neither trait nor method--all the correlations in the first column of the lower left hand block except the correlation inside the method diagonal (.54). Comparing the .88 validity index with the four relevant correlations in the column below it (.56, .58, and .44), we find that .88 is higher than all of these correlations, providing evidence of discriminant validity. Considering all the comparisons which bear on this hypothesis, we find that 23 of 40 comparisons for speaking and 38 of 40 comparisons for reading support the first hypothesis of discrimination.

The second discriminant validity hypothesis is that the validity indices should be higher than correlations between measures of different traits measured by the same method. To test this hypothesis, we compare the set of validity indices with the correlations in the method diagonal. For example, discriminant validity is demonstrated by comparing validity index .88 with the .54 correlation between the interview test of speaking and the interview test of reading. Discrimination is not demonstrated, however, when we compare the validity index .51 with the .54 correlation. Considering all the comparisons which bear on this hypothesis, we find that discriminant validity is supported in 7 of 10 comparisons for speaking and 4 of 10 for reading.

The effect of method is particularly noticeable in measures using translation or self-rating methods. Of the correlations in the method diagonal, the intercorrelations between measures 3, 4 and 5, which employ translation and self-rating methods (.64, .69, and .68) are higher than those between measures 1 and 2 (.54 and .46), which do not use these methods.

### Confirmatory Factor Analysis

Confirmatory factor analysis, as noted above, is a technique for statistically evaluating the goodness of fit of competitive causal models to a body of data. Over ten causal models were tested against our data, each involving different explicit assumptions regarding trait-method interaction. The two models with the best fit make three explicit assumptions: 1) no trait-method interaction, 2) some method interaction, and 3) equal factor loadings across raters. One of these models posits three method factors and one trait factor (the unitary trait model), while the other model posits three method factors and two trait factors (the divisible trait model). A comparison of these models is given in Table 3.

To test the hypothesis of distinct speaking and reading traits, we examine the difference between the chi squares of the unitary trait model ( $\chi^2 = 50.722$ ) and of the divisible trait model ( $\chi^2 = 34.980$ ). The difference between these chi squares (15.742,  $df = 1$ ) is significant at  $p < .001$ , and we therefore reject the hypothesis that a single language trait underlies the measures. The correlation between the two trait factors (.524), however, suggests the possibility of a single factor common to these two trait factors.

Having found a model which provides a good fit for the data, we can examine the loadings of each measure on each of the five factors, to determine the relative importance of trait and method in each. Factor loadings and uniquenesses for the ten measures are given in Table 4. The high loading of the oral interview measures on the speaking factor (.819), compared to the relatively lower loading of the oral translation measures (.563) and the oral self-rating measure (.298), indicates that the oral

interview method provides a better measure of speaking ability than do the translation and self-rating methods. An examination of the loadings of the interview, translation and self-rating measures on the reading factor leads us, by similar reasoning, to conclude that of the three methods examined, the translation measure (with a loading of .756 on the reading factor) provides the best measure of reading ability.

TABLE 3  
Comparison of chi squares for two models

1 Trait (Unitary)	$\chi^2 = 50.722$ df = 30 p = .0104
2 Traits (Divisible)	$\chi^2 = 34.980$ df = 29 p = .2052 $r_{\lambda_{ti} \lambda_{tj}} = .524$
$\chi^2$ difference	$\chi^2_2 - \chi^2_4 = 15.742$ df = 1 p < .001

TABLE 4  
Factor loadings (and standard errors) for ten measures

<u>Measures</u>	<u>Speaking</u>	<u>Reading</u>	<u>Interview</u>	<u>Translation</u>	<u>Self-Rating</u>	<u>Uniqueness</u>
Oral Interview 1	.819 (.082)	.000	.459 (.126)	.000	.000	.113
Oral Interview 2	.819 (.082)	.000	.459 (.126)	.000	.000	.132
Oral Translation 1	.568 (.091)	.000	.000	.729 (.098)	.000	.175
Oral Translation 2	.568 (.091)	.000	.000	.729 (.098)	.000	.137
Oral Self-rating	.298 (.097)	.000	.000	.000	.734 (.108)	.357
Reading Interview 1	.000	.155 (.140)	.972 (.085)	.000	.000	.034
Reading Interview 2	.000	.155 (.140)	.972 (.085)	.000	.000	.017
Reading Translation 1	.000	.756 (.097)	.000	.611 (.133)	.000	.044
Reading Translation 2	.000	.756 (.097)	.000	.611 (.133)	.000	.070
Reading, Self-rating	.000	.216 (.113)	.000	.000	.834 (.104)	.235

Loadings of the measures on the three method factors are consistent with these conclusions. Specifically, the oral measures load less heavily on the interview method factor (.459) than on the translation method factor (.729) and on the self-rating method factor (.734). This indicates that the effect of method on oral test scores is least for the interview method. In other words, assuming we were interested in maximizing the effect of the trait we are trying to measure and minimizing the effect of method, we would choose the interview method to measure oral ability.

Looking at the effect of method on the reading test scores, we find that the translation method (which loads .611 on the reading tests) affects the reading test scores less than the self-rating method (.834) or the interview method (.972). We conclude, therefore, that of the three methods used in the study, the one which minimizes the effect of test method on the reading test scores is the translation method.

#### CONCLUSION

This study has yielded two types of results, methodological and empirical. With respect to methodology, we believe that the application of confirmatory factor analysis to multitrait-multimethod data enables us to identify and quantify the effects of trait and method on measures of language proficiency, and provides us a far clearer picture of the nature of this proficiency than has been available with other types of analysis.

With respect to our empirical findings, we feel we have found evidence demonstrating the convergent and discriminant validity of the FSI interview. Of the three methods used to measure proficiency in speaking, the interview method evidenced the largest trait component and the smallest method component. In addition, we feel we have found strong evidence for the distinctness of speaking and reading as traits and thus support for the divisible trait hypothesis. Subsequent research will examine (1) the extent to which a common factor may underly these two distinct traits, and (2) the composition of these individual traits.

*FOOTNOTES*

<sup>1</sup>This research was partially supported by a grant from the Research Board, University of Illinois, and by the authors' departments, the Division of English as a Second Language, University of Illinois, Urbana-Champaign and the Department of English, University of Utah. We also express our appreciation to the participants in the colloquia on the construct validation of oral language tests, held at the 1979 and 1980 TESOL national conventions, for their contributions to every phase of the research described in this paper. We would also like to thank Robert L. Linn, of the Department of Educational Psychology, University of Illinois, for his counsel on confirmatory factor analysis. Any shortcomings in our application of this procedure are, of course, our own.

<sup>2</sup>We express our appreciation to Pardee Lowe and Ray Clifford, who provided us with an excellent four day intensive training program in the administration and rating of oral interviews at the Language School, Central Intelligence Agency.

## REFERENCES

- Adams, M. L. and J. R. Frith. 1979. *Testing kit: French and Spanish*. Washington, D.C.: Foreign Service Institute, United States Department of State.
- Alwin, Duane F. 1974. Analyzing the multitrait-multimethod matrix. In H. L. Coster, ed. *Sociological methodology 1973 - 1974*. San Francisco: Jossey-Bass.
- Althausser, Robert B. 1974. Inferring validity from the multitrait-multimatrix: another assessment. In H. L. Coster, ed. *Sociological methodology 1973 - 1974*. San Francisco: Jossey-Bass.
- Althausser, R. P., T. A. Heberlein and R. A. Scott. 1971. A causal assessment of validity: the augmented multitrait-multimethod matrix, in H. M. Blalock, ed. *Causal models in the social sciences*. Chicago: Aldine - Atherton.
- American Psychological Association. 1974. *Standards for educational and psychological tests and manuals*. Washington: American Psychological Association.
- Brütsch, S. M. 1979. Convergent/discriminant validation of prospective teacher proficiency in oral and written French by means of the MLA cooperative language proficiency tests, French direct proficiency tests for teachers (TOP and TWP), and self-ratings. Unpublished Ph.D. dissertation, University of Minnesota.
- Campbell, D. T. and D. W. Fiske. 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin* 56, 2.
- Clark, J. L. D. 1975. Theoretical and technical considerations in oral proficiency testing. In R. L. Jones and B. Spolsky, eds. *Testing language proficiency*. Arlington, VA: Center for Applied Linguistics.
- Clark, J. L. D. 1978. *Direct testing of speaking proficiency: theory and application*. Princeton: Educational Testing Service.
- Clark, J. L. D. 1979. Direct vs. semi-direct tests of speaking ability, in E. Brière and F. B. Kinofotis, eds. *Concepts in language testing: some recent studies*. Washington, D.C.: Teachers of English to Speakers of Other Languages.
- Clifford, Ray T. 1978. Reliability and validity of language aspects contributing to oral proficiency of prospective teachers of German. In J. L. D. Clark, ed., *Direct testing of speaking proficiency: theory and application*. Princeton: Educational Testing Service.
- Clifford, Ray T. 1980. Convergent and discriminant validation of integrated and unitary language skills; the need for a research model. In A. S. Palmer and P. J. M. Groot (eds.) *The validation of oral*

- proficiency tests: an introduction.* Washington: Teachers of English to Speakers of Other Languages.
- Corrigan, A. and J. A. Upshur. 1978. Test method and linguistic factors in foreign language tests. Paper presented at the 1978 TESOL convention, Mexico City.
- Cronbach, L. J. 1971. Test validation. In R. L. Thorndike, ed. *Educational measurement*, 2nd Ed. Washington: American Council on Education.
- Cronbach, L. J. and P. E. Meehl. 1955. Construct validity in psychological tests. *Psychological Bulletin* 52, 4.
- Jöreskog, K. G. 1969. A general approach to confirmatory maximum likelihood factor analysis, *Psychometrika* 34.
- Jöreskog, K. G. and O. Sörbom. 1978. LISREL: User's guide to analysis of linear structural relationships by the method of maximum likelihood, Version IV, Release 2. Chicago: National Educational Resources, Inc.
- Lowe, Pardee, Jr. 1976. *Handbook of question types and their use in LLC oral proficiency tests.* Washington, D.C.: CIA Language School.
- Lowe, P. and R. T. Clifford. 1980. Developing an indirect measure of overall oral proficiency (ROPE). Paper presented at the colloquium on validation of oral proficiency tests, 1980 TESOL Convention, San Francisco.
- Oller, J. W. and K. Perkins (eds.) 1980. *Research in language testing.* Rowley, Mass.: Newbury House.
- Palmer, A. S. and P. J. M. Groot, eds. 1980. *The validation of oral proficiency tests: an introduction.* Washington: Teachers of English to Speakers of Other Languages.
- Specht, D. A. 1976. SPSS: statistical package for the social sciences version 6 user's guide to subprogram reliability and repeated measures analysis of variance. Ames: Statistical Laboratory, Iowa State University.
- Stevenson, D. K. 1980. Beyond faith and face validity: the multitrait-multimethod matrix and the convergent and discriminant validity of oral proficiency tests. In A. S. Palmer and P. J. M. Groot, eds. *The validation of oral proficiency tests: an introduction.* Washington: Teachers of English to Speakers of Other Languages.



THE ROLE OF TEACHER FEEDBACK IN PREVENTING  
THE FOSSILIZED ERRORS OF SECOND LANGUAGE LEARNERS<sup>1</sup>

H. Douglas Brown

The key to successful prevention of fossilized errors of second language learners lies in the kind of feedback learners receive in the process of communication. In operant conditioning models of learning, the effectiveness of feedback is seen as a factor of the degree of reinforcement of certain forms of language. A more recent model of feedback (Vigil and Oller 1976) implies that the best feedback is that which is an optimal blend of "cognitive" and "affective" feedback. In what Brown (1980a) described as the "optimal distance model," a further dimension of feedback is discovered. Practical implications for teachers are discussed.

One of the most frustrating problems that face second language teachers in the classroom is the problem of the persistence of errors in the production of second language learners despite repeated efforts to help the learners overcome those errors. The term fossilization has been used to describe the tendency of second language learners to manifest certain production errors well into fluent stages of language acquisition. While we seem to be able to identify and even to diagnose the sources of fossilized errors, we are usually at a loss to know how to treat the condition. But we do know something about the prevention of fossilization, and that is the subject of my comments here.

The key to successful language learning lies essentially in the feedback that the learner receives. In order to examine the role of feedback in preventing fossilization, I will first summarize some theoretical constructs that can help us to understand the phenomenon, and then discuss the implications of those constructs for practical contexts in the language classroom. Three theoretical viewpoints will be brought to bear on the problem: (1) reinforcement theory as explicated by B. F. Skinner; (2) a model of feedback proposed by Vigil and Oller (1976) (3) a notion of the relationship of stages of acculturation to second language learning, designated as the optimal distance model (Brown 1979).

## Reinforcement Theory

Reinforcement theory is no stranger to language teachers. Prevailing methodologies of the 1950s and 1960s were built solidly on principles of reinforcement theory. Skinnerian reinforcement theory, or operant conditioning, differed markedly from earlier classical (stimulus-response) conditioning. In the typical classical model, stimuli were related to responses such that the stimuli were viewed as the key to an organism's achievement of criterion behavior. Skinner, on the other hand, stressed the crucial importance of the response, and particularly emphasized the essential role of the reinforcement of desired responses for the ultimate learning of behavior. Skinner contended, and showed under numerous experimental conditions, that the positive reinforcement of "emitted" responses -- that is, responses which were manifested without the direct, intervening administration of external stimuli -- served to produce better long-term learning than "elicited" responses. Skinner was therefore not really interested in the study of stimuli. In human behavior most stimuli in natural contexts are difficult to observe anyway, and require mentalistic speculation for adequate identification, which of course is sheer heresy for a hard-nosed behaviorist!

From pigeons to human beings, Skinner showed the strength of the operant conditioning model. Rewards for desired, emitted behavior (operants) produce effective results. Punishment, or the application of aversive stimuli, has only a short-term effectiveness. A dog will more readily learn to move to one side of a box when a light is flashed on if it is rewarded with food than if it is punished for not moving fast enough by an electric shock received through a grid in the floor of the box. Similarly, a child will better learn not to throw baseballs in the house by being positively rewarded for throwing them outside than by being punished for throwing them inside.

Skinner's operant conditioning model is not incompatible with what some consider to be a conflicting theoretical stance, namely, cognitive psychology. Cognitive psychologists stress the importance of the independent, creative organism acting upon the environment, as opposed to the environment (reinforcer) acting upon the learner. But the effectiveness

of positive reinforcement depends upon the attention and perception of the learner. If the learner does not discern reinforcement with his own powers of perception, there will actually be no reinforcement!

In second language learning we can identify a learner's spoken (or written) utterance as a response,<sup>2</sup> and the hearer's verbal or nonverbal feedback as a reinforcer. The degree to which the learner perceives and processes the feedback will determine the effectiveness of the reinforcement and the strength of the learned behavior.

### A Model of Feedback

The verbal and nonverbal behavior which teachers manifest in response to -- or in reinforcement of -- second language learners' utterances can be termed feedback. Vigil and Oller (1976) outlined a model of feedback for second language learning that helps us to refine our understanding of the nature of reinforcement in the process of second language learning. Vigil and Oller defined two levels of feedback in linguistic interchange. Affective feedback is a verbal or nonverbal message which says to the speaker: "I value you, I want to understand you, I want communication to take place," or, conversely, "I don't value you, I don't want to understand you." Cognitive feedback conveys a message which indicates that "I understand your message, what you said is clear," or "I don't understand your utterance, it is not clear." Both levels of feedback can be either positive or negative, as indicated above, though it is obvious that positive and negative feedback represent poles on a continuum with innumerable "shades of gray" in between. It is also apparent that the differences between affective and cognitive feedback are not entirely clear-cut: a degree of overlap is present in any linguistic interchange.

Vigil and Oller maintained that for communication to take place at all, positive affective feedback is essential. If there is negative affective feedback, the result is that the speaker will -- at least ultimately -- abort the attempt to communicate. Even neutral affective feedback (the lack of any reinforcement whatever) might serve to extinguish the desire of the speaker to produce a linguistic utterance (or operant). On the cognitive level of feedback, however, positive and negative feedback

can both serve to produce desired behavior. Positive cognitive feedback ("I understand you") has the effect of reinforcing or rewarding the production response emitted by the learner. Negative or neutral cognitive feedback (a message that indicates the hearer has not understood production or some portion of production), when accompanied by positive affective feedback, serves to lead the speaker to rephrase or in some way attempt another response. Negative cognitive feedback is commonly associated with attempts by teachers to correct students' errors in some way.

The most interesting implication of Vigil and Oller's model is that cognitive feedback must be optimal in order to be effective. Too much negative cognitive feedback -- a barrage of interruptions, corrections, and overt attention to malformations in the form of a speaker's utterance -- often leads the learner to shut off his attempts at communication. He perceives that so much is wrong with what he is producing that there is little hope for much to be right with it. On the other hand, too much positive cognitive feedback -- willingness of the hearer-teacher to let errors go uncorrected, to indicate understanding when understanding may not really be present -- serves to reinforce the errors of the speaker-learner, and the result is the persistence, or fossilization, of such errors. The task of the teacher is to discern the optimal tension between positive and negative cognitive feedback: providing enough positive cognitive feedback to encourage continued communication, but not so much that errors are reinforced, and providing enough negative cognitive feedback to call attention to crucial errors, but not so much that the learner is discouraged from attempting to speak in the second language.

### Optimal Distance Model

A third theoretical stance to be directed to an understanding of feedback is the one that I have called the optimal distance model (Brown 1980a). Generally speaking, cultural anthropologists recognize a number of different stages which persons go through in the process of acculturation: the first, a stage of euphoria and excitement in the new culture; second, a stage of "culture shock" -- depression, bitterness, despair, and

anger characterize the person attempting to adjust to the new and strange culture where nothing seems to square with the world view of the native culture; the third stage is characterized by gradual but faltering recovery from culture shock; the fourth stage is one of either assimilation or adaptation in the new culture.

Lambert (1967) offered the first clue connecting acculturation with second language learning. Lambert found that certain learners of French as a second language experienced feelings of anomie, or homelessness (characteristic of the end of the second stage where a person feels that he belongs to neither culture) just as they began to break through to fluency in the second language. That is, the beginning of language mastery coincided with the end of the second stage of acculturation and the onset of the third stage. Lambert's findings agree with recent research by Acton (1979) in suggesting that there is a point in the acculturation process at which language mastery may be optimal. That optimal point occurs when the learner senses that his own social or psychological distance between the native and target cultures is relatively equal; he perceives that he is in cultural limbo, neither too close nor too far from either culture. The reason for language mastery at this point is the presence of appropriate tension for the learner. With the apron strings of the native culture almost fully severed, the learner has a high degree of motivation to survive in the culture, and language is seen as the major tool for survival. The learner therefore seeks positive reinforcement of emitted linguistic responses; and that reinforcement (which affirms the ability to "think" in the second culture as well as to communicate) leads to linguistic mastery.

The model suggests that there is a "critical period" for second language acquisition within the second culture<sup>3</sup> such that before the end of stage two the learner lacks sufficient perception of reinforcement for his linguistic responses, and that at the end of stage two all necessary reinforcing agents are present and perceived in the environment of the learner. It may also be the case that after the end of the second stage, if someone has learned to cope psychologically in the second culture without linguistic mastery, the learner fails to perceive the need for

correctness and the repeated reinforcement of incorrect responses leads to the fossilization of errors and to less than successful mastery of correct forms of the second language (see Brown 1980b: 138-139).

### Some Practical Implications

I have summarized three theoretical constructs which have a bearing on our understanding of the importance of feedback in the process of learning a second language. The Skinnerian operant conditioning model really serves to undergird and support Vigil and Oller's conception of the various levels of feedback which teachers can give. Feedback is the reinforcement of verbal operants that will lead to the establishment of those operants as desired and appropriate behavior. In the optimal distance model the learner's perception of reinforcement in the linguistic environment will be colored by his psychological state of mind; that state reaches an optimum at a particular cultural stage.

How can the language teacher find practical implications in these constructs for classroom teaching? First, the basic (Skinnerian) model tells us that the most important, if not exclusively important, behaviors in the classroom are those reinforcing responses that a teacher gives to the verbal operants emitted by the learners. Everything that a teacher says and does reinforces something. Those reinforcers, as they accumulate through days and weeks and months of classroom instruction, become the crucial determiners of the long-term success of the second language teacher. They are more crucial than all the best-laid plans of textbooks and other prepared material. While Skinner de-emphasizes the role of the stimulus in the control of behavior, he in no way rules out the important influence of good materials (see, for example, Skinner 1968). Prepared material, in the Skinnerian sense, serves not so much to goad or prod but through careful design to reinforce the learner along scientifically constructed pathways to success. But in language learning, the best learning takes place in the very act of free, meaningful communication, where there is little that can be predicted scientifically. The teacher must therefore be prepared, for every operant which a learner emits, to give positive affective feedback, and to discern the optimal form of cognitive

feedback. Until the learner is fully adapted or assimilated into the second culture, the teacher needs to be particularly sensitive to cross-cultural variations in the learner's perception of reinforcers. How do you know that a learner from another culture is perceiving your feedback as reinforcing? Often you do not know. The best recourse is for the teacher to invoke a breadth of knowledge and sensitivity to cultural differences in perception.

The second general implication of the three constructs under consideration here lies in the affective nature of second language learning. Self-esteem, empathy, inhibition, alienation -- just a few of the many affective factors influencing second language acquisition -- have been explored in previous research. The feedback which teachers provide always possesses such affective attributes. Some time ago, a group of high school students from the French club at the high school came to my door selling candy to support their club. Delighted with the apparent enthusiasm of this group of young language learners, I broke into my own French: "C'est très bien, alors, vous parlez un petit peu de français! Entrez s'il vous plait." Nonplussed at this unexpected response, the students quizzically looked at me, then at each other, and without saying a word, embarrassingly left my front doorstep, never to be seen again! I never found out whether they really were from the French club, though I expect they were. My response unintentionally turned out to be threatening and alienating instead of reinforcing. Our affective feedback in the classroom must always be supportive, empathic, and non-defensive.

Third, in considering the administration of positive reinforcement, the teacher is faced with an age-old dilemma: to correct or not to correct. A strict interpretation of the operant conditioning model tells us that perhaps no correction is warranted -- we should only offer positive reinforcement of desired behavior and ignore the undesired. Such a Skinnerian approach may be adequate as a general guideline, but we have learned from other models and from experience in language teaching that neutral feedback on an error can be perceived as a positive reinforcement of the error. These other models give us some guidelines on correction.

(1) The principle of affective feedback tells us to allow the learner to

express himself as freely as possible without interruption. A learner of English as a second language was once asked to introduce himself on the first day of class. The learner stood up, and in fear and trembling produced an obviously carefully devised sentence: "Ladies and gentlemen, allow me to introduce myself and tell you about the 'headlights' of my past." At that point a teacher clearly should not interrupt to correct the learner! (2) The principle of optimal cognitive feedback tells us to help the learner to avoid miscommunication through certain errors. When a native Spanish speaker in a university setting asked his teacher and classmates to come to his party in his "department," the teacher wisely ascertained the correct item, "apartment," thus avoiding the awkward possibility of a group of would-be party-goers in the university's Department of Civil Engineering. (3) The optimal distance model tells us that many corrections -- however crucial they appear to the teacher -- may go unnoticed by the learner if his acculturation timetable is either well before or well after the sociocultural critical period of learning. When that critical period has apparently been achieved by the learner, the teacher might at that point capitalize on the learner's linguistic readiness. (4) Finally, a decision about which form of correction (direct correction, subtle indication of an error, expansion, and grammatical explanation) will be most effective can sometimes be made on the basis of the theoretical constructs elaborated upon here. A correction which casually expands the learner's utterance will usually give more positive affective feedback than direct correction or grammatical explanation. On the other hand, when a learner is hopelessly "tangled up" in a sentence to the point that he cannot extract himself, a direct correction -- a rescue from the teacher -- may then be more reinforcing. It is the teacher's task to assess all the social, psychological, cultural, and linguistic attributes of a learner's discourse and then to provide appropriate forms of correction.

A fourth and final implication of our theoretical models goes beyond correction to a possibility that merges cognitive feedback with affective feedback: compliments, verbal and nonverbal, offered in reward of correct, appropriate responses. While most language teachers are generous with



positive reinforcement, the genuineness of the teacher's "that's good!" is sometimes questionable in the humdrum activity of the language class. Correct responses are sometimes the result of the very careful attention and hard work of the learner. The teacher as the provider of the reinforcement of such operants needs to offer genuine rewards.

### Conclusion

An interesting and perhaps a significant principle emerges from this consideration of theoretical constructs in interpreting the role of feedback in second language acquisition. In virtually every variable affecting second language acquisition there is a principle of tension at work, a principle in which optimum mid-points between poles hold the key to successful acquisition. In the operant conditioning model the tension between reinforcement and response will determine the degree of learning. There is a point of tension between positive and negative cognitive feedback at which reinforcement is sufficient. There is tension in the process of acculturation that gives rise to a critical period of learning: optimal learning will occur when there is optimal distance between the native and target cultures. These and other constructs, in which the tension between two points is crucial, lead us to conclude that there is no singularly significant variable leading to second language mastery. The successful performance of each learner will be a factor of the complex of variables that define the learner, the language itself, the context, and the goals of the learner. Theoretical models provide a series of transparencies, as it were, with which the classroom teacher can view each individual learning situation. The effectiveness of teacher feedback in the second language classroom will depend upon the successful perception of points of tension in the theoretical transparencies and then upon appropriate practical application.

## FOOTNOTES

<sup>1</sup>A revised and condensed version of a lecture delivered at TESOL conferences in 1979.

<sup>2</sup>It is of course true that comprehension (aural or reading) also comprises a set of "responses." Comprehension responses may be more "elicited" than "emitted." Moreover, comprehension responses are less "publicly observable" and therefore subject to more mentalistic speculation than a Skinnerian would care to indulge in. I will therefore limit my comments here to discussion of production responses.

<sup>3</sup>Note that I am restricting the application of the optimal distance model to second language learning within the context of the second culture. The acquisition of a second language for specialized purposes, learning a second language in the native culture (e.g., English in India), and foreign language learning (e.g., French in the United States) must be excluded from such application.

## REFERENCES

- Acton, William. 1979. Perception of lexical connotation: Professed attitude and socio-cultural distance in second language learning. Unpublished doctoral dissertation, University of Michigan.
- Brown, H. Douglas. 1980a. The optimal distance model of second language acquisition. *TESOL Quarterly* 14:2. In press.
- Brown, H. Douglas. 1980b. *Principles of Language Learning and Teaching*. Englewood Cliffs, NJ: Prentice-Hall.
- Lambert, Wallace. 1967. A social psychology of bilingualism. *The Journal of Social Issues* 23:91-109.
- Vigil, Neddy, and John Oller. 1976. Rule fossilization: A tentative model. *Language Learning* 26:281-295.



PSYCHOMETRIC AND EDMETRIC APPROACHES TO LANGUAGE TESTING:  
IMPLICATIONS AND APPLICATIONS

Gary A. Cziko

Recently, the basic principles underlying educational measurement have undergone a thorough re-examination resulting in a distinction between what are called psychometric and edumetric approaches to testing. Generally defined, a psychometric test is one that is designed to maximize inter-individual differences resulting in scores that can be interpreted only by comparison with the scores of an appropriate norm group. In contrast, an edumetric test is designed to maximize intra-individual differences (i.e., learning) and to yield scores that are meaningful in themselves without reference to the performance of other examinees. The realization of this distinction has given rise to a lively debate concerning the relative strengths and weaknesses of the two types of tests and their proper use in various testing situations. However, this important distinction appears to have gone largely unnoticed in the field of language testing, where the controversy concerning discrete-point and integrative testing techniques has attracted most attention. This paper examines the implications and applications of the psychometric-edumetric distinction for language testing and its interaction with the discrete-integrative distinction. It is suggested that language tests possessing both psychometric and edumetric characteristics as well as purely edumetric tests can be created by adapting existing integrative testing techniques. Four such adaptations are examined involving the use of auditory-auditory, auditory-visual, visual-auditory, and visual-visual modalities for presentation and response, respectively. Plans for future research on language testing are also presented. 1

INTRODUCTION

Recently, the basic principles underlying educational measurement have undergone a thorough re-examination. The result of this re-examination has been an important distinction between what are called psychometric and edumetric properties of educational tests. Generally defined, a psychometric test is a test which has been primarily designed to maximize individual differences on the variable being measured, resulting in scores that can be interpreted only in terms of comparing them with the scores of an appropriate comparison group on the same test. In contrast, an edumetric test is designed primarily to yield scores which are meaningful

without reference to the performance of others. Although the reader may not be familiar with the use of the terms psychometric and edumetric, most readers will be familiar with two examples of psychometric and edumetric tests, namely, norm-referenced and criterion-referenced tests, respectively. The awareness of this basic distinction in the properties of educational tests has given rise to a lively controversy regarding the relative strengths and weaknesses of psychometric and edumetric measurement and their proper use in various testing situations (see Block 1971; Ebel 1971).

There have also been recent important developments in techniques used to measure language proficiency. These developments are primarily due to a re-examination of what it means to know a language and new methodologies for teaching languages. The most important change in language testing techniques has been the introduction and use of integrative tests of language proficiency (e.g., cloze and dictation tests) in contrast to the more traditional though still popular discrete-point tests (e.g., vocabulary and grammar tests; see Oller 1976). It is interesting to note that in spite of the many books and papers published on language testing within the last several years, little mention has been made of the important distinction between the psychometric and edumetric approaches to testing. It may be that the changes which have taken place in linguistic theory (from structural to transformational to pragmatic views of language competence) have caused researchers interested in language testing to be more concerned with what it is they are trying to measure than the underlying statistical assumptions and properties of various kinds of tests. However, it appears that a consideration of the psychometric and edumetric dimension in language testing and the interaction of this dimension with the discrete point-integrative one is long overdue and may be useful in shedding new light on basic issues in the measurement of language proficiency.

The remainder of this paper consists of five sections. The first section outlines the essential differences between the psychometric and edumetric approaches to educational measurement and contrasts norm-referenced with domain-referenced and criterion-referenced tests. The

second section discusses the implications of the psychometric-edumetric distinction for language testing and examines the two approaches to measurement in terms of their compatibility with current theories of language learning and language use. A review of discrete-point and integrative approaches to language testing is presented in the third section. The fourth section considers the interaction of the psychometric-edumetric and discrete point-integrative dimensions with respect to language testing and suggests new testing procedures and modifications of existing integrative tests to create measures having both psychometric and edumetric characteristics. Plans for continued research on language testing are offered in the final section.

#### PSYCHOMETRIC AND EDUMETRIC APPROACHES TO TESTING

Until fairly recently, the techniques used in the development of educational tests were borrowed directly from the psychometric tradition in psychology. These techniques, developed to measure psychological constructs such as intelligence, are based on two important criteria: reliability and validity. Reliability is usually considered to be the extent to which a measuring instrument is free of random measurement error. A totally reliable test of English vocabulary, for example, would always assign the same score to a particular examinee, assuming that his knowledge of English vocabulary had not changed from one testing session to the next. Validity, however, has to do with the extent to which a measuring instrument measures what it is designed to measure. Hence, a test of academic aptitude designed to measure one's aptitude for success in college is a valid test to the extent with which it accurately predicts college achievement. While the reliability of an instrument does not depend on its validity, the validity of a test is limited by its reliability.

Although it is obvious that the criteria of reliability and validity are important for any type of measuring instrument, the classical psychometric procedures for establishing the reliability and validity of a test are based on the dispersion of test scores. One index of reliability is based on the covariance of test scores on the same test or

parallel tests given at different times. Validity is usually defined in relation to the proportion of test score variance which predicts or explains individual differences on performance on an appropriate criterion. A test of academic aptitude would therefore have high reliability if the same test on different occasions or parallel forms of the test ranked each individual of a group examinees in the same way each time. The test would have high validity if each examinee was more successful academically than all examinees with lower test scores but less successful academically than all examinees with higher test scores. This psychometric view of reliability and validity based on the variance of test scores has led test developers to construct tests that maximize individual differences in order to obtain high indices of reliability and validity. This maximization of individual differences is typically obtained by including items which on the average are answered correctly by only half of the examinees since items that are either passed or failed by nearly all of the examinees contribute very little to the variance of the total test scores. It should be noted that this approach is consistent with the idiographic tradition in psychology which postulates that individual differences exist and the purpose of any measuring instrument is to capture and magnify these differences.

Since tests developed according to psychometric principles yield scores which maximize individual differences, these scores in raw form are usually meaningless. To say that Maria got 40 items right on a test of English vocabulary consisting of 100 items does not provide much useful information on Maria's knowledge of English vocabulary. It is for this reason that individual scores derived from psychometric tests are usually compared with the scores of an appropriate comparison or "norm" group. If we know now that only 30% of the students in Maria's class obtained scores lower than 40, we have a better notion of Maria's knowledge of vocabulary, although this may also mean that Maria's classmates are all native speakers of English whereas Maria is not. Since the use of a norm group in the interpretation of psychometric test scores used for educational measurement is essential, these tests are referred to as norm-referenced tests and the raw scores obtained are usually converted



to some sort of standard score (e.g., stanines, percentiles, z-, or T-scores) which compare the performance of an individual with that of the norm group. To recapitulate, the two essential characteristics of norm-referenced tests are that they are designed to maximize the variation among test scores and that test scores can only be interpreted in comparison with an appropriate norm group.

Recently, however, measurement specialists have begun to question the appropriateness of norm-referenced tests in educational settings. Popham (1975:128,129) attributes this re-examination of norm-referenced tests to the popularity of programmed instruction in the 1950's which was based on the belief that using the appropriate teaching methodology, virtually all learners could master the essential objectives of virtually any course of study. This implied that effective instruction would result in relatively little variation in individuals' study of the course's goals. Since norm-referenced tests are designed to maximize individual variance and since the items of a norm-referenced test are typically not directly related to the objectives of any particular instructional program, it became clear that other methods were needed to assess the effectiveness of various educational programs and teaching methodologies. This led to the realization of the importance of what Carver (1974) referred to as the edumetric dimension of tests and to the consequent definition, development, and use of educational tests which are primarily edumetric in nature, i.e., domain-referenced and criterion-referenced tests.

It has already been noted that a test with edumetric properties yields scores that are meaningful without reference to the performance of others. This is possible since edumetric tests normally include only items which test a carefully defined domain of behaviors. For example, should we wish to measure a fourth-grade pupil's knowledge of spelling, we could create a test that includes all 200 words from the list of fourth-grade spelling words specified in the curriculum. If the pupil spelled 150 of these words correctly it would not be necessary to compare his performance to the performance of others for interpretation since this score in itself indicates that the pupil is able to correctly spell 75% of the words for which he is responsible. Of course, it may not be feasible to administer a test of

such length and so we may decide to construct a test of only 20 words randomly sampled from the 200-word list. In this case a score of 15 words correctly spelled would allow us to infer or estimate that the pupil can spell correctly 75% of the words on the full list. We are now interpreting a test score not in reference to a norm group but rather in reference to a well-defined domain of behaviors. A test which allows such an interpretation is referred to as a domain-referenced test.

However, we may feel that 75% simply isn't good enough. We may decide that being able to spell correctly 30% of the sixth-grade words is the criterion for success in spelling and that the pupil must continue to work on his spelling until this criterion is achieved in order to proceed to the next unit of instruction. Since the pupil's score is now being interpreted in reference to a specified criterion in addition to a domain of behaviors, we will refer to the test as a criterion-referenced test. Although norm-referenced tests scores may be also interpreted with reference to a criterion (e.g., a medical school may accept only those students who score among the top 10% of all students taking the Medical College Admission Test), a criterion-referenced test for the purpose of this paper will be defined as a domain-referenced test for which some criterion has been set. This is in keeping with Denham's (1975) view of norm-, domain-, and criterion-referenced measurement.<sup>2</sup> As Denham notes, we may also make norm-referenced interpretations of scores obtained on domain- or criterion-referenced tests. For example, we may wish to compare the above pupil's spelling score of 15 out of 20 with the performance of others in his class and find that his score is higher than only 35% of his classmates. While such an interpretation is possible, it should be noted that the items for the test were chosen to be representative of a defined behavioral domain and are not necessarily the best items for discriminating between good and poor spellers. However, while a domain- or criterion-referenced test may give some useful norm-referenced information, a test which is constructed for the sole purpose of magnifying individual differences is normally not appropriate for estimating an examinee's performance on tasks other than those included in the test.

What about the reliability and validity of edumetric tests? Since

edumetric tests are not constructed to maximize score variance, it is generally agreed that the classic psychometric techniques of deriving coefficients of reliability and validity using correlational techniques are not appropriate for edumetric tests. There appears to be little consensus, however, on exactly how one should go about to assess the reliability and validity of edumetric tests. Carver (1974) has recommended that the validity and reliability of an edumetric test be based on the test's sensitivity to educational growth (learning) and the consistency with which it measures this growth. Popham (1975), however, makes no mention of growth and argues that if an edumetric test is properly constructed reliability should not be a problem since the generation of items from a well-defined domain should result in a homogeneous set of items and a test with high internal consistency and reliability (p. 152). With respect to validity, Popham states that the sine qua non of edumetric measurement is descriptive validity, i.e., "the degree to which a domain definition adequately delimits the nature of a set of test items and, further, the degree to which the test items are congruent with the domain definition" (p. 159) and that "it is difficult to conceive of many situations in which an educational evaluator will need measures that do more than adequately describe the performance of learners" (p. 155). This is an interesting thought and much in contrast to the psychometric notions of concurrent or predictive validity which require the use of an appropriate criterion measure to establish the validity of a test. Popham's view of validity has important implications for language testing which will be discussed in the following section.

Since a domain of behaviors must be well defined before one can construct a domain- or criterion-referenced test, the reader may well wonder how this is done for testing knowledge of areas other than fourth-grade spelling words. The problem of domain definition has received considerable attention from Hively and his associates (Hively, Maxwell, Rabehl, Sension, & Lundin 1973) although their work deals primarily with the areas of mathematics and science, two subjects which appear to be particularly well suited to domain specification. Even in science and mathematics, however, there are still no universally accepted procedures for defining domains or

for generating items. This contrasts with the discipline of psychometric or norm-referenced measurement where there exist generally accepted procedures for test construction and item selection (see for example, Magnusson 1968). The problems of edumetric test construction appear even more acute in the social sciences, humanities, and language arts. The specific problems encountered in the edumetric measurement of language proficiency will be considered in the next section.

#### IMPLICATIONS OF THE PSYCHOMETRIC-EDUMETRIC DISTINCTION FOR LANGUAGE TESTING

Now that the essential characteristics of the psychometric-edumetric distinction have been described and norm-, domain-, and criterion-referenced measurement defined, the next issue to consider is the implications of these measurement approaches for language testing.

Looking at the testing of language proficiency first from a psychometric viewpoint, one may ask if it makes sense to consider language proficiency along a quantitative dimension with high proficiency at one end, low proficiency at the other, and varying degrees of proficiency between the two. This notion does seem to be in keeping with most notions of second-language proficiency since some second-language learners are clearly more proficient than others and levels of proficiency may range all the way from zero competence to native-speaker competence in the second language. This notion of a wide range of proficiency levels is also compatible with current theoretical views of second-language learning which see the learner passing through a series of language systems (or interlanguages) each one more similar than the preceding one to the system of linguistic knowledge possessed by the native speaker (see Selinker 1974). It appears, therefore, that it does make sense to be concerned with individual differences in second-language proficiency and that in certain situations we may well wish to develop and use tests which have maximum sensitivity to these differences.

However, the fact that psychometric norm-referenced test scores are not meaningful until compared to an appropriate second-language norm group is considerably less appealing in the measurement of second-

language proficiency. Since in measuring language proficiency we have at least one natural reference point to use as a criterion (i.e., the native speaker), it would appear preferable in many situations to be able to determine the distance separating the language learner from native-speaker proficiency than to say that he scored better than a certain percentage of his classmates. This type of criterion-referenced interpretation is usually not possible with norm-referenced tests of language proficiency since the items chosen for such tests are those which maximize the test score variance of second-language learners and are not chosen from a domain of items which test essential components of language skills possessed by native speakers. It is therefore not possible to interpret a score of 75 out of 100 on a norm-referenced test of second-language proficiency as representing 75% of the knowledge of a native speaker since there is no assurance that a native speaker would respond correctly to all 100 items or that items are a representative sample of relevant language behaviors.

An edumetric criterion-referenced test of second-language proficiency would be necessary to yield scores directly interpretable in terms of a performance standard (e.g., native-speaker proficiency) and there have been attempts to construct such tests (see Cartier 1963; Defense Language Institute 1975). Such a test if properly constructed would appear to be quite attractive since if the domain of relevant behaviors has been adequately defined and if the test items are representative of this domain, there would appear to be no need to validate the test against an external criterion in the way that is necessary for a norm-referenced test. This is an especially desirable feature due to the difficulty of finding adequate criteria of language performance. The principle difficulty with the construction of such a test is specifying the domain of behaviors from which test items will be selected. Although methods of task analysis and content analysis have been devised to establish relevant domains of language behavior (see Defense Language Institute 1963:30-38) it appears that these methods are best suited to highly specific second-language situations such as the behaviors required for success in special purpose language courses or in very specific

occupational settings. It would seem virtually impossible to specify all of the language behaviors which might be required of a language user in less specific situations. Even if these behaviors could be defined, however, there remains the problem of translating these behaviors into items or tasks that can be practically used in testing situations. For instance, suppose that we have determined that an essential required behavior of a group of second-language learners is being able to properly greet persons of both sexes of various social classes at various times of day. If we devise a series of paper and pencil items to test this behavior it is clear that the behavior thus elicited will be quite different from the verbal behavior required in a real greeting situation. The same holds true for domains developed from a more grammatical viewpoint. Although we may consider proper use of verb tenses to be part of the domain of required behaviors, performance on test items developed to assess knowledge of verb tense may not necessarily be predictive of real language use, especially if we have used discrete-point items which have not been validated against real language performance. Finally, even if we could specify the domains of interest and properly translate these behaviors into test items, it seems unlikely that we could ever say that functioning in a given language situation requires so much of ability A, so much of ability B, etc. It seems more likely that language abilities interact in complex and perhaps compensatory ways so that weakness in a particular skill (e.g., vocabulary knowledge) may be made up for by strength in another (e.g., sensitivity to contextual constraints).

It appears then that while the psychometric norm-referenced approach to language testing may be valid and useful in certain situations, this approach does not yield test scores which are directly interpretable in terms of the useful performance standard of native speakers or other defined groups. While the edumetric criterion-referenced approach to the measurement of language proficiency would theoretically permit such an interpretation, there appear to be serious conceptual and methodological problems in the construction

of such tests. Before considering possible solutions to these problems, it will be useful to examine in some detail the distinctions between the discrete-point and integrative approaches to language testing.

#### DISCRETE-POINT AND INTEGRATIVE APPROACHES OF LANGUAGE TESTING

Oller (1976:275,276) describes a discrete-point approach to language testing as one which "requires the isolation of skills (such as listening, speaking, reading, and writing), aspects of skills (such as recognition versus production, or auditory versus visual processing), components of skills (such as phonology, morphology, syntax, and lexicon), and finally, discrete elements (such as phonemes, morphemes, phrase structures, etc.)" whereas the integrative approach "tries to measure global proficiency and pays little attention to particular skills, aspects, components, or specific elements of skills." This distinction forms the basis for what is undoubtedly the most well known and controversial issue in language testing and it is interesting to note that although there appears to be a trend toward increasing use of integrative tests (e.g., cloze and dictation), virtually all commercially available tests of language competence continue to comprise subtests of language skills which are primarily discrete-point in nature.

Much has been written about discrete-point vs. integrative debate in language testing and the reader is referred to Oller (1976) and Davies (1979) for interesting and contrastive treatments of this issue. A question of primary concern to this paper is one which has not been treated in the literature, i.e., the relation between the psychometric-edumetric dimension and the integrative-discrete point dimension of language tests. At first glance it would appear that discrete-point tests are primarily psychometric while integrative tests are more edumetric in nature, although closer scrutiny clearly shows that this is not necessarily the case. Discrete-point tests appear more in keeping with the psychometric tradition since the use of dis-

crete, unrelated items allows for item analysis and the subsequent retention of those items that contribute most to test score variance and there is no doubt that the primary purpose of most if not all discrete-point tests of language proficiency is to maximize individual differences. In fact Spokslly (see Valette 1977:308,309) has used the term "psychometric-structuralist" to describe the discrete-point approach to language testing advocated by Lado (1961) and his followers, implying that the testing of knowledge of discrete points of language structure and the psychometric approach to testing go (or at least went) hand in hand. It is true that Lado appears to have been greatly influenced by the psychometric tradition in testing, devoting entire chapters to fairly classical psychometric definitions of validity, reliability, and item analysis. However, there are parts of Lado's book that do not conform to the psychometric approach to testing. In discussing item selection, for example, Lado (p. 346) mentions "editing on the basis of performance by native speakers" and suggests eliminating items missed by 10% or more of native speakers. Lado also states that "even if an item does not correlate with the total test score of the students we are justified in keeping the item as part of the test provided the problem...is part of the skill we wish to test..." (p. 349) and "As the content of the test we can and usually must select a sample of the things that have to be learned. This sample should be randomly selected" (p. 20). Both of these suggestions appear quite edumetric in nature, first suggesting a sort of criterion-referenced approach (using the performance of native speakers as a criterion) and then recommending a domain-referenced approach where an item's contribution to total test score variance is not as important as prior definition of the behaviors to be included in the test and the use of this domain in the random selection of test items.

Integrative language tests appear to be more in keeping with the edumetric approach to testing. This seems to be the case for two main reasons. First integrative tests do not readily lend themselves to item analysis and subsequent deletion of items which do not dis-



criminate well between examinees with high and low language proficiency. In fact, for a dictation, translation, or oral interview it is not exactly clear what constitutes an item. This is less difficult for cloze tests although even here it is not normally possible to delete an item without rendering the test meaningless. Second, most integrative tests (e.g., cloze, dictation, translation) are based on a text which in many ways can be regarded as a sample of language chosen from all the possible relevant texts that could have been used for the test. Of course, makers of cloze texts do not normally put all possible texts in a large barrel and then pick one out at random. Nonetheless, if the text chosen can be considered representative of the type of language the examinee is likely to be required to deal with, it can be considered a sample in the domain-referenced sense of the word. In fact the only reason anyone ever draws a random sample of anything is to be sure that in the long run the elements chosen are representative of the universe of which they form a part. It is for this reason that we often speak of a representative sample and if we can be convinced that a sample is representative we care little about how it was actually chosen.

While one then may tend to see integrative tests as more edumetric than psychometric in nature for the above two reasons, it seems that such tests are typically put to psychometric use. Carroll (1972) who first advocated the use of integrative tests of language proficiency stated that "an ideal English language proficiency test should make it possible to differentiate, to the greatest possible extent, levels of performance which are relevant to the kinds of situations in which the examinees will find themselves" (p. 319) and goes on to stress the importance of the predictive validity of such tests, both very psychometric concerns. However, Carroll does add that "Ideally, one should have a list of all possible items which one might cover, and draw a sample by random sampling techniques" and that it is important to "define as carefully as possible the total area from which one is sampling" (p. 320) quite in contrast to the psychometric concerns mentioned by him on the previous page. Oller, the person who

has done more than anyone to research, explain, and promote the use of integrative language tests, seems to emphasize their psychometric use in stating that "it is the variance in test scores, not the mean of a certain group or the score of a particular subject on a particular task that is the main issue" (1979:272).

It is remarkable to observe that both Lado and Carroll, early proponents of quite different approaches to language testing in terms of the discrete point-integrative dimension, saw a need for both psychometric and edumetric approaches to language testing and in fact were concerned with domain definition, item generation, and item sampling even before these terms became part of the measurement specialist's jargon. However, neither Lado nor Carroll explained how a language test, whether discrete point or integrative, could be a random sampling of a specific domain of language tasks and yet be able to maximally differentiate low- and high-proficiency examinees.

#### FOUR TYPES OF LANGUAGE TESTS

Since it appears that the psychometric-edumetric and discrete point-integrative dimensions are independent (i.e., both discrete-point and integrative tests of language proficiency can be constructed according to either psychometric or edumetric principles), we will now examine the four combinations of these two dimensions.

First, let us consider a psychometric discrete-point test of language proficiency. Such a test would be composed of items dealing with individual language elements and selected to maximally differentiate examinees possessing different levels of knowledge of the elements tested. An example would be a test of vocabulary, grammar, or preposition usage composed of separate multiple-choice items having close to a .50 average level of difficulty. In addition, psychometric definitions of reliability and validity would require that the items be highly inter-correlated and that total test scores correlate highly with an appropriate criterion.

What would scores on such a test mean? Since the test items were selected to maximize total score variance, it is clear that the

interpretation of any individual test scores will necessitate comparison to an appropriate norm group. What the test actually measures, however, will depend on its demonstrated validity. For example, if a test of vocabulary knowledge has been shown to correlate highly with future grades in language courses, the test could be considered a test of language aptitude. However, a discrete-point psychometric test of language proficiency that has not been validated against an appropriate criterion is of little use, even if the test is reliable and has face validity (i.e., appears to measure what it is supposed to measure). This is because the items of a discrete-point test typically involve tasks that bear little resemblance to tasks of actual language use. Carroll has noted that "if we limit ourselves to testing only one point at a time, more time is ordinarily allowed for reflection than would occur in a normal communication situation..." (1961:318). Also, discrete-point tests usually provide little in the way of contextual information. For example, although an examinee may not be able to determine the meaning of a word presented in isolation on a discrete-point test of vocabulary, he may be able to determine meaning when presented in a meaningful context of a reading passage, a conversation, or a course lecture. In addition, even if an examinee is able to formally state a grammatical rule or recognize a correct or incorrect application of a rule, he may still fail to use the rule in his own production (see Seliger 1979). For these reasons, a discrete-point, psychometric test of language proficiency must be validated using an appropriate norm group.

Once such a test has been validated, however, it does not seem at first glance to matter that a discrete-point test may require behaviors that are unlike those required in real communicative settings or that the results of tests which purport to measure different skills are intercorrelated. It appears quite likely that performance on a discrete-point test may correlate highly with communicative language behavior simply because the learning of explicit language rules and the development of communicative competence in a language are usually correlated with the length of exposure to the language. For example,

a foreign student enrolled in an American university and attending a course in English as a second language will gain explicit knowledge of language rules (the kind of knowledge that will help the student's performance on discrete-point tests) from instruction in the language course as well as communicative competence and an implicit knowledge of the rules of the language from using the language communicatively both inside and outside the classroom. Because of this indirect relationship, discrete-point test scores may correlate with language proficiency not because one needs to have explicit knowledge of discrete bits of language rules to communicate but because discrete-point performance and language proficiency are indirectly related. It should be noted, however, that while a discrete-point language test may be valid for one group or type of learners, it may not be valid for other groups. It may be that a discrete-point test of grammar predicts ability to use the language in communicative settings but only for those second-language learners who have had formal instruction in the language. This same test may not be a valid measure of language proficiency for those learners who have learned the language in informal, non-classroom settings and it has been shown that certain discrete-point items are in fact more difficult for native speakers than second-language learners of the language (Angoff and Sharon 1971). Thus, one could argue that a discrete-point test must be validated anew for any group of learners which differs in any substantial way from the group originally tested. This appears to be a serious problem with discrete-point tests which must be kept in mind in spite of the evidence presented by Farhardy (1979) which suggests that with certain groups of L<sub>2</sub> learners there may be no important statistical differences between discrete-point and integrative tests.

The second type of test to discuss is the edumetric discrete-point test. Such a test would also be composed of discrete-point items but the criteria for selecting and keeping items would not be based on psychometric principles. While items for a psychometric discrete-point test would be chosen to maximize total test score variance, maximize reliability and correlate highly with a meaningful

criterion measure, items for an edumetric discrete-point test would be chosen for quite different reasons. In fact, there are two quite distinct approaches that one could use to select items for an edumetric discrete-point test of language proficiency.

The first is the domain-referenced approach of sampling items from a domain that has been used by Cartier (1963). Some problems with using this approach for language testing have already been mentioned and Oller raises yet another in describing the difficulty of sampling from a universe which is "infinitely large and non-repetitive" (1979: 184). Even if one could sample from this universe, however, one could only generalize the performance on the test to a universe of discrete-point items. This may be useful if one's use of language is restricted to classroom drills and discrete-point tests but not if one is interested in language use in real communicative settings. Of course scores on such a test may predict performance on other language criteria but since items have not been chosen to represent a real communicative domain or to maximize test score variance it seems unlikely that these scores would permit accurate and useful predictions.

The second way of selecting items for an edumetric discrete-point test would be to use Carver's (1974) original notions of edumetric test validity. According to Carver, an edumetric test should be composed of items having maximum sensitivity to learning. That is, items should be chosen to maximize intra-individual differences before and after the learning experience. Thus, an ideal edumetric test of knowledge learned during a French course would be one on which most students score zero before the course and most successful students score 100% at the end of the course. While such a test would be by definition very sensitive to learning taking place during the course, it would not be very useful in telling us how much French was actually being learned compared to other students in other courses or if what was being learned would be of any use outside of the classroom. A more promising approach would be to use items which discriminate between the pre- and post-instruction scores of language learners who can be considered successful by some external criterion. Then we

would know that a person with a high score has not only learned something during the course but that what he has learned was also learned by previous successful students. This approach to edumetric discrete-point language testing could have many uses, although it does not seem to have been ever used. Possible applications of this technique as well as adaptations of it will be discussed later.

We now move on to integrative language tests and we will examine these first from a psychometric and then from an edumetric viewpoint. Seen from a psychometric viewpoint, integrative tests appear to be somewhat of a puzzle. Although the two most common forms of integrative tests, cloze and dictation, are not constructed according to traditional psychometric principles of test construction, they nevertheless appear to have many of the essential characteristics of psychometric tests. That is, although these tests have not usually been subject to item analysis and checks of reliability and validity, they nonetheless seem to be very sensitive to inter-individual differences in language proficiency and as a result correlate quite highly with many psychometric discrete-point tests of language proficiency (Oller, 1976).

Another indication of the fine sensitivity of integrative tests to inter-individual differences in language proficiency comes from the evaluations of French immersion programs in Montréal. One particular study (Cziko, Holobow, and Lambert 1977) involved the comparison of the French language proficiency of four different groups of pupils at Grades 4 and 5. The first group was made up of a class of native English-speaking children whose instruction in French was limited to approximately 40 to 50 minutes of French-as-a-second-language instruction per day. The second group were English-speaking children who had had the same type of French instruction as the first group during Grades 1 through 3 but during Grade 4 had received almost all their classroom instruction in French by a native French-speaking teacher (i.e., had one year of what is referred to as French immersion). The third group were also English-speaking children but this group had had four to five years of French immersion experience

(although both French and English had been used as media of instruction during Grades 2 through 5). Finally, the fourth group comprised native French-speaking children attending French-language schools. The results of a French cloze test administered to these four groups revealed significant differences between all possible pairs of the four groups--Group 4 scoring significantly higher than Group 3, Group 3 significantly higher than Group 2, and Group 2 in turn scoring significantly higher than Group 1. The ability of the cloze test to discriminate among all four levels of French-language proficiency is even more notable when one considers that none of the other French-language tests used in the evaluation (tests of writing, reading comprehension, and speaking) consistently discriminated among all four groups, although these latter tests required considerably more time to construct.

Why is it, then, that integrative tests are such good psychometric tests even though they are not constructed according to the psychometric principles of test construction? There are at least two possible reasons. Part of the answer probably lies in the nature of the behavior required by all integrative tests. As Oller (1979) has noted, integrative tests (or pragmatic tests as he has defined them) require the same kind of behavior that is required in actual meaningful language behavior. This is because integrative tests require the examinee to use his "grammar of expectancy," i.e., the ability to use knowledge of the syntactic, semantic, discourse, sociolinguistic, and extralinguistic rules of language behavior to formulate what he is about to say or write as well as to predict what he is about to hear or read within a limited period of time. Therefore, an integrative test would appear to be a much more direct test of language proficiency than a discrete-point test which requires behaviors very different from anything a language user is ever required to do in a real communicative setting. (It should be noted that the actual overt behavior required by a cloze or dictation test is also not something that is normally done in real language situations, but the underlying cognitive processes are thought to be essentially the same.) So, if

integrative tests are in fact more direct measures of language proficiency and if we assume that second-language proficiency can vary widely from zero competence to native-speaker competence, it follows that integrative tests would be quite sensitive to inter-individual difference in language proficiency even though they are not constructed along the lines of traditional psychometric tests.

The other part of the answer probably lies in the type of response format used by integrative tests. A test can only have good psychometric qualities if the items it comprises are able to discriminate between low and high-ability examinees. In fact, the primary purpose of item analysis as it is applied to most psychometric discrete-point tests is to find which items do not discriminate between high and low scorers so that these items may be revised or deleted. However, since discrete-point tests are typically composed of multiple-choice questions, it is possible that a low scorer will pass an item simply by chance (call this a gift) and that a high scorer will fail an item that is actually within his competence due to a cleverly disguised distractor (call this a gyp). Whenever a gift or a gyp occurs, the true variance of the test scores is reduced and the sensitivity of the test to individual differences (the sine qua non of a psychometric test) is reduced. Gifts and gyps are less likely to occur on integrative tests, however. Gifts are rare since integrative tests are usually not presented in multiple-choice formats and so if the examinee has absolutely no idea of what the sixth word of the dictation is or what could fit the third cloze test blank, he is very unlikely to come up with the correct word by chance. Now, one may argue that performance on a cloze test is nothing but a series of guesses and this is true; however, these guesses are likely to result in a correct response only if they are based on an adequate knowledge of the language which is exactly what the test is designed to measure. Gyps also appear to be less of a problem on integrative tests since there are no cleverly written distractors to tempt the unwary examinee. The only distractors present are those formulated by the examinee himself and are most likely the result of an incomplete or deviant



knowledge of the language. However, gyps may be quite common on cloze tests scored according to the exact-word criterion since the examinee may often come up with a perfectly acceptable response which is nonetheless considered incorrect. This probably partly explains why cloze tests scored according to the acceptable-word criterion (which greatly reduced the probability of gyps) have been found to be more sensitive to differences in language proficiency than cloze tests scored using the exact-word method (see Alderson 1979; Oller 1972).

To summarize, integrative tests have proven to be sensitive psychometric measures of language proficiency in spite of the fact that they are not constructed according to traditional psychometric principles. Their psychometric qualities are likely due to the fact that they are direct measures of language proficiency and that gifts and gyps are unlikely to occur.

Let us now take a look at integrative tests from an edumetric viewpoint. We have already seen that integrative tests are consistent with one notion of what constitutes an edumetric test since they are based on a sample of language. However, if we consider the other usual criteria of edumetric tests we see that integrative tests as they are used today do not fit these criteria. First of all, the raw scores provided by integrative tests are not usually meaningful in themselves and there seems to have been no attempt to develop criterion-referenced integrative tests in spite of the fact that in most testing situations we can identify a group of language users (native speakers or successful second-language learners) whose performance could be used as a criterion. Even if this is done, however, we still have a problem since if our criterion groups scores 75%, an examinee scoring 75% may have attained this score with a very different pattern of passed and failed items which would seem to indicate that in spite of his score reaching criterion, his proficiency is not the same as that of the criterion group. Second, if we accept the notion that edumetric tests should maximize intra-individual differences (i.e., learning) there appears to be no research demonstrating that in fact integrative tests do this. Fortunately, although integrative tests as they now

exist do not seem to have many edumetric qualities, it appears that they could be transformed into tests with excellent edumetric characteristics with relatively minor changes.

The trick is to design a test so that an examinee's score is directly interpretable with respect to a criterion and yet is very sensitive to intra-individual differences, something which fortunately appears quite easy to do. For example, imagine we are interested in measuring how close a group of foreign university students come to native English speakers in their ability to handle the style of English that news broadcasters use in Illinois. Our first task is to find a group of native speakers that we feel is representative of the criterion group to which we wish to compare our examinees and a text that we judge to be representative of Illinois broadcast English. We must then decide what input-output modalities we wish to use. There are four choices: auditory-auditory, auditory-visual, visual-visual, and visual-auditory for input and output, respectively. We opt for the auditory-visual combination and so shall use a test of dictation. We now administer the chosen text as a dictation test to each of the members of our criterion group, adjusting the number of presentations, the rate of presentation, and the size of the chunks presented so that a typical criterion individual's score is very close to, but not quite, 100%. The dictation is tape-recorded and we now have a integrative language test which has been calibrated against a meaningful criterion and which should yield scores that are directly interpretable (without comparison to the scores of the other second-language examinees) indicating how close an examinee is in language proficiency to the criterion group.

Note that we have also created a test that would appear to be very sensitive to learning. According to Carver (1974), an edumetric test sensitive to learning should assign a score of zero to examinees with no knowledge of what the test is designed to measure and assign a score of 100% to examinees with criterion-level knowledge of what is being measured. Since examinees with no knowledge of the language being tested are sure to score zero (remember that integrative tests

give no gifts) and examinees with criterion-group knowledge of the language should score close to 100% (since the test has been calibrated using a criterion group), we have apparently succeeded in constructing an integrative test of language with all desirable edumetric properties. It should also be noted that we have done nothing which might compromise the demonstrated psychometric properties of the dictation test. In fact, by calibrating the dictation test using an appropriate criterion group it would seem that we have actually enhanced the psychometric properties of the test for examinees with language proficiency levels below that of the criterion group. This is because we have positioned the ceiling of the test at a level where individual differences are not of interest while scores below this level are free to vary. This contrast with an uncalibrated dictation test which may be either too easy or too difficult for many of the examinees resulting in a restricted range of scores that do not reflect inter-individual differences.

Of course, this is only one example since edumetric integrative tests can be varied along at least four different variables: the input-output modalities (task), the criterion group, the language type, and the scoring procedure. The first variable is the choice of the input-output modalities. Four possibilities have been mentioned, although only two of these (i.e., auditory-visual in the form of dictation and visual-visual in the form of cloze) appear to be presently in use. The auditory-auditory combination in the form of elicited imitation would seem to be the natural choice for measuring language proficiency without involving reading or writing (e.g., for young children or illiterates) and has been recently recommended by Hameyer (1978) and Oller (1979:289-295) as a useful and valid approach to language testing.

The visual-visual combination is the one used for cloze tests. However, cloze tests do not seem to be very well suited to edumetric measurement since it appears to be difficult to construct a cloze test on which a reference group (even highly educated native speakers whose performance is scored using the appropriate-word method) will score close to 100%. A more promising approach using the visual-visual

combination would be similar to dictation except that the input would be visually presented language. This could be simply done in a classroom by projecting slides with chunks of written text for specific durations or putting the entire text on an overhead transparency and using masks to present chunks of the text (presumably after letting the examinees read the entire text first in the same way that an entire dictation is usually presented before allowing the examinees to respond). Like dictation, the size of the chunks and the duration of each presentation would be determined using an appropriate reference group.

Finally, an adaptation of this technique could be used to create a visual-auditory test. This test would be similar to the preceding visual-visual test except that the examinee would orally repeat what he had just been visually presented. Unfortunately, both this technique and elicited imitation share the disadvantage of requiring that the output be tape-recorded which precludes group administration unless a language laboratory with recording facilities can be used.

The choice of task will depend on the purpose of the test. For our foreign university students the auditory-visual combination would probably be most appropriate if we were interested in the ability of the students to follow lectures and take notes although if reading ability were considered more important we might opt for the visual-visual test. As previously mentioned, the auditory-auditory combination is the only possibility for testing young children and illiterates. Finally, the visual-auditory test seems most appropriate when we wish to measure reading without involving writing.

The second variable is the choice of a criterion group. While in the previous example we used native speakers, there are other possibilities. For example, if we are interested in determining whether foreign students have sufficient command of English in order to be successful in studying at an American university we may well want to use as our criterion group foreign students at an American university who we know have been successful in their studies. The criterion group will again depend on the purpose of the test and could con-

ceivably range all the way from highly educated native speakers to beginning second-language learners who have been judged successful according to some criterion.

The third variable refers to the different types of language that can be used in edumetric integrative tests. The possibilities are limitless in spite of the fact that language tests have traditionally been restricted to fairly formal registers of standard written and spoken language. As always, the choice will depend on the purpose of the test. For example, our foreign students should be tested using a representative sample of the type of language they are likely to encounter in course lectures. One way to do this would be to use part of an actual lecture that had been edited to remove ungrammatical sentences and slips of the tongue. Similarly, reading passages representative of "textbook" English could be used for visual-visual and visual-auditory tests.

But we are certainly not limited to formal or standard varieties of language. For example, if part of a student's reason for studying French is to be able to converse with working-class Parisians, an auditory-auditory or auditory-visual task using text representative of a conversation between two Parisian blue-collar workers should give the student (and teacher) valuable information about his ability to do so. The advantage of using integrative tests for such purposes is that performance on such tests would appear to depend not only on the examinee's knowledge of vocabulary, syntax, and semantics of the target language but also on the ability to make use of discourse constraints and knowledge of the rules of language use in specific situations (see Hymes's 1972 notion of communicative competence). Obviously, the language material used for such a test, whether it be taken from written or oral sources, must be coherent text which conforms to the discourse and sociolinguistic constraints of the language. Although the auditory-auditory modality combination has been referred to as elicited imitation, the common practice of using a set of unrelated sentences for elicited imitation (e.g., Naiman, 1974) would not be appropriate for use in such a test.

The fourth variable is the scoring procedure and can be of essentially two types: verbatim scoring or scoring for comprehension. Oller (1979:293-295) explains these two procedures and discusses how they could be used to investigate language comprehension, language preference, and bidialectalism.

#### PLANS FOR RESEARCH ON LANGUAGE TESTING

It should be clear by now that the author considers the four above-mentioned adaptations of integrative tests as having great potential for the measurement of all aspects of language proficiency. This is because such tests appear to be quite direct measures of language proficiency, have both psychometric and edumetric properties, and yield meaningful scores which are directly interpretable with reference to a criterion group. However, since such tests have apparently not yet been constructed and put to use, it remains to be demonstrated empirically that they do have the desirable characteristics which they appear to possess a priori and are useful in a wide variety of testing situations. Basic research needs to be done to investigate the effects of the four above-mentioned variables (i.e., task, criterion group, language and scoring) on performance on such tests. Another particularly interesting area of research is to examine the sensitivity of such measures to knowledge of the rules of language use (often referred to as the sociolinguistic or ethnographic rules of speaking a particular language). It is now widely recognized that such knowledge is crucial to the language learner's ability to communicate in real social settings although the definition of what constitutes the elements of such knowledge has proved elusive and its measurement problematic (however, see Walters 1979). It will also be of interest to determine the sensitivity of the proposed tests to knowledge of different language dialects, the usefulness of such tests in determining the language dominance and preference of multilinguals, for the placement of foreign university students, and for setting entrance and exit criteria for bilingual education programs.

*FOOTNOTES*

<sup>1</sup>Much of this paper was written while the author was a post-doctoral fellow at the Faculté des sciences de l'éducation, Université de Montréal. The author was financially supported during this time by a post-doctoral fellowship granted by the Ministère de l'éducation du Québec. Requests for reprints and other correspondence should be sent to Gary A. Cziko, Department of Educational Psychology, University of Illinois, Urbana, Illinois 61801.

<sup>2</sup>It should be noted that although this interpretation of the distinction between domain-referenced and criterion-referenced tests is consistent with Denham's (1975) view, it is perhaps not the most widely accepted (cf. Glass 1978; Popham 1978). It is used here because it is consistent with the concept of criterion as it is used later in this paper.

## REFERENCES

- Alderson, J. C. 1979. The cloze procedure and proficiency in English as a foreign language. *TESOL Quarterly* 13, 219-227.
- Angoff, W. H., and Sharon, A. T. 1971. Comparison of scores earned by native American college students and foreign applicants to U.S. colleges. *TESOL Quarterly* 5, 129-136.
- Block, J. H. 1971. Criterion-referenced measurements: Potential. *School Review* 79, 289-298.
- Carroll, J. B. 1971. Fundamental considerations in testing for English language proficiency of foreign students. In H. B. Allen and R. M. Campbell (eds.) *Teaching English as a second language* (2nd ed.). New York: McGraw-Hill.
- Cartier, F. A. 1968. Criterion-referenced testing of language skills. *TESOL Quarterly* 2, 27-38.
- Carver, R. P. 1974. Two dimensions of tests: Psychometric and edumetric. *American Psychologist* 28, 512-518.
- Cziko, G. A., Holobow, N. E., and Lambert, W. E. 1977. *A comparison of three elementary school alternatives for learning French: Children at Grades 4 and 5*. Unpublished manuscript, McGill University, Department of Psychology.
- Davies, A. Language testing (2 parts). 1978. *Language Teaching and Linguistics Abstracts* 11, 148-158; 215-231.
- Defense Language Institute. 1975. *Handbook for conducting task analyses and developing criterion-referenced tests of language skills*. Lackland, Texas: English Language Center, Lackland Air Force Base.
- Denham, C. H. 1975. Criterion-referenced, domain-referenced and norm-referenced measurement: A parallax view. *Educational Technology* 15, 12:9-12.
- Ebel, R. L. 1971. Criterion-referenced measurements: Limitations. *School Review* 79, 282-283.
- Farhardy, H. 1979. The disjunctive fallacy between discrete point and integrative tests. *TESOL Quarterly* 13, 347-358.
- Glass, G. V. 1978. Standards and criteria. *Journal of Educational Measurement* 15, 237-261.
- Hameyer, K. 1978. *Testing oral proficiency via elicited imitation*. Paper presented at the International Congress of Applied Linguistics, Montreal.
- Hively, W., Maxwell, G., Rabehl, G. J., Sension, D. B., and Lundin, S. 1973. *Domain-referenced curriculum evaluation*. Los Angeles: Center for the Study of Evaluation, University of California.



- Hymes, D. On communicative competence. 1972. In J. B. Pride and J. Holmes (eds). *Sociolinguistics*. Harmondsworth, England: Penguin.
- Lado, R. 1961. *Language testing*. London: Longman.
- Magnusson, D. 1968. *Test theory*. Reading, Mass.: Addison-Wesley.
- Naiman, N. 1974. The use of elicited imitation in second language acquisition. *Working Papers on Bilingualism* 2, 1-37.
- Oller, J. W., Jr. 1972. Scoring methods and difficulty levels for cloze tests of ESL proficiency. *Modern Language Journal* 56, 151-158.
- Oller, J. W., Jr. 1976. Language testing. In R. Wardaugh and H. D. Brown (eds.). *A survey of applied linguistics*. Ann Arbor, Mich.: University of Michigan Press.
- Oller, J. W., Jr. 1979. *Language tests at school*. London: Longman.
- Popham, W. J. 1975. *Educational evaluation*. Englewood Cliffs, N.J.: Prentice-Hall.
- Popham, W. J. 1978. *Criterion-referenced measurement*. Englewood Cliffs, N.J.: Prentice-Hall.
- Seliger, H. W. 1979. On the nature and function of language rules in language teaching. *TESOL Quarterly* 13, 359-370.
- Selinker, L. Interlanguage. 1974. In J. S. Schumann and N. Stenson (eds.). *New frontiers in second language learning*. Rowley, Mass.: Newbury House.
- Valette, R. M. 1977. *Modern language testing* (2nd ed.). New York: Harcourt Brace Jovanovich.
- Walters, J. 1979. *Language variation, politeness and bilingual children*. Paper presented at the meeting of the American Educational Research Association, San Francisco.



A PEDAGOGICAL INTERPRETATION OF GENERATIVE PHONOLOGY

I. THEORETICAL FOUNDATIONS

Wayne B. Dickerson

This paper is the first in a series of articles on the practical applications of generative phonological research. The fundamentals of taxonomic and generative phonology are discussed, and the implications of each for teaching pronunciation are drawn out. The two theories are compared with respect to the linguistic phenomena they describe and the capabilities they attribute to the learner. The comparisons show why the generative model is the more adequate theoretical foundation for a course in English pronunciation. An outline of behavioral objectives based on this model is provided. Concluding the study is a progress report on the status of applied generative phonology in the profession today.

INTRODUCTION

During the last thirty years, the letters TESL, which were originally a simple acronym, have become the symbol of an internationally recognized professional and academic endeavor. Not only have the size, diversity and status of our field grown over the years, but also, happily, our understanding of language structure, our appreciation of the learning task, and our repertoire of teaching methodologies have increased with time.

Our improved understanding of language, language learning and language teaching has left its mark on every area of ESL teaching. But the teaching of pronunciation has been less visibly affected by advances in knowledge than other aspects of the ESL curriculum. ESL pronunciation texts covering the vowels, consonants, stress and intonation of English, whether published in 1950 or 1980, differ little from each other in their language content, in their concept of the learner, or in their behavioral objectives. Their remarkable uniformity derives from an adherence to taxonomic phonology, a model of the sound system which has been obsolete in technical circles for more than fifteen years. In short, while the field of TESL, as a whole, has changed in response to new insights, pronunciation instruction has not been equally

sensitive to new developments.

Given the richness of phonological insight now available to us from technical sources, we might ask: Why has pronunciation teaching not kept pace with linguistic developments? What can be done to incorporate these developments into our instructional materials? To answer these two questions, we must understand the influence of theoretical models on pedagogical decisions. One purpose of this paper, then, is to compare the old linguistic model of taxonomic phonology with the new model of transformational generative phonology, to draw out the pedagogical implications of each, and to formulate a well-motivated approach to the challenge of pronunciation teaching. A second purpose of this paper is to introduce a series of reports on eight years of research aimed at implementing the generative-based approach outlined here.

#### ABSTRACT AND CONCRETE PRONUNCIATION GOALS

The pedagogical differences between a taxonomic-based text and a generative-based text arise from the important theoretical issues separating the two positions. Let us begin by stating the goals of pronunciation teaching in terms which are generally acceptable to proponents of both theories and then identify the areas in which there is the greatest disagreement.

A general definition of teaching goals is possible because, for all their differences, taxonomic and generative phonologists agree on a fundamental fact about language, namely, that part of it is abstract and part of it is concrete. Since this bipartite nature of language figures in our definition of teaching goals, the distinction merits some elaboration.

Speakers communicate with each other by means of sounds. However, they are able to communicate only because they share an understanding of how the sounds carry meaning. Their shared understanding is an abstract phenomenon, the study of which is called *phonology*. Their use of speech sounds is a concrete phenomenon, the study of which is called *phonetics*. Both phenomena together, the mental organization of

sounds and the physical sounds themselves, are language, and both are prerequisite to oral-aural communication.

Reflected here is a time-honored distinction between the abstract language system and the concrete use of the system. At the turn of the century, Ferdinand de Saussure, the father of modern linguistics, called the system, *la langue*, and the manifestation of the system in an actual utterance, *la parole*. More recently, Noam Chomsky added a psychological dimension to the distinction and referred to the speaker-hearer's knowledge of language structure as his *competence* and his use of the structure to speak, hear, read and write as his *performance*.

At root, there is a single dichotomy here, whether it is stated in specific terms as phonology-phonetics, or in general terms as langue-parole or competence-performance. Either way, this dichotomy not only rallies linguists of all persuasions but also allows us to conceptualize the goals of teaching pronunciation.

On the one hand, we want our students to be able to behave as English communicators, that is, to sound as though they are speaking English when they are, and to distinguish the significant sounds of English when they hear them. They must learn to shape phonetic articulations and sharpen phonetic recognition. *We want to promote concrete performance skills in the area of phonetics.* On the other hand, we want our students to be more than good mimics and sound discriminators. We also want them to use sounds to communicate. They must learn how to organize sounds into patterns to convey meaning. That is, *we also want to promote an abstract native-like competence in the area of phonology.*

The goals described above have been defined with few specifics and within a framework which underlies most linguistic research. At this level of generality, pedagogically-minded taxonomists and generativists will find little dispute with each other. But when we discuss the content of what is to be learned, serious differences surface. There is least disagreement about the content of the phonetics instruction. Because this is the more concrete level, both schools agree in large measure on the vowel and consonant phones of English,

the levels of stress and the kinds of pitch phenomena distributed over words and utterances. There are differences, but they are relatively superficial. By contrast, there is great disagreement about the content of phonological instruction. Here the differences are fundamental, because they go to the heart of issues which separate the two theoretical positions.

In the interest of narrowing our discussion, we shall set aside consideration of the first goal--imparting phonetic skills--so that we can focus more adequately on the second goal--imparting phonological understanding. The next section then, concerns the contrasting theoretical models of phonology and their implications for teaching pronunciation.

#### QUERIES AND THEORIES

What is the structure of the sound system? What kind of organization regulates the use of sounds in language? These questions deal in the main with the content of phonology, the point at which the taxonomist and the generativist differ the most. The reason for their differences is that phonological analyses are shaped by what the linguist sets out to explain and by the research assumptions he adopts. So, if we want to understand the content of a phonology, we must first understand the phonologist's point of view: What questions does he seek to answer? What limitations does he put on himself? Let us address these matters before considering the phonological models themselves.

#### Queries

The taxonomist asks: How have native speakers organized the vast phonetic detail of their speech so that it communicates meaning? His approach to this question is guided by a commitment to a strict empirical research strategy. He begins with objective data--phonetic data. From the phonetic level, he must describe each subsequent level--phonology, morphology, syntax--exhaustively in terms of what is known to that point. Thus, he works from the concrete to the abstract, and

at no level of analysis does he make use of information from a higher level. He rejects any dependence on intuition, either his own or his language helper's. He wants his procedure to be so mechanical that any investigator could arrive independently at the same analysis given the same raw data.

The generativist's query focuses on certain native speaker abilities. What kind of system enables the speaker [1] to code meaning in sound (a question similar to the taxonomist's), [2] to predict the proper word stress, vowel quality and consonants for unfamiliar words, and [3] to know that sounds are related to each other at different levels of the grammar? Illustrating this third ability, the language user knows that the [eyt] and [ət] renderings of the word *postulate* are connected at the level of syntax--verb vs. noun, and that the [t] and [č] of *depart*, *departure* are connected at the level of morphology. To explain such abilities, the generativist uses a principled mentalism in his research. He is rigorously rational: Every element in his analysis must be supported by a well-motivated explanation, and every proposal must be considered an hypothesis subject to disconfirmation. Thus, he is able to work from the abstract levels toward the concrete level. Beginning with syntax, he makes all structural information available in the phonological analysis. As data, he is able to use the speaker's intuitions, and he also depends heavily on his own resourcefulness to understand how the system works. His aim is not replicability of findings but a fully explicit account of the speaker's competence.

### Theories

The taxonomic and generative explanatory models of the native speaker's langue/competence are as different as the research interests which they address. The structure of these models can be compared beginning with what each theorist considers the raw material--the input--for the phonological analysis.

In the taxonomic analysis,

phonetics supplies the raw material, described as

objectively as possible... This objective description of speech is what phonemics [phonology] takes and studies in order to see what generalizations can be made about its structure. (Francis 1958:120)

So, the taxonomist begins with a detailed analysis of concrete speech sounds--*phones* ([ ]). At the next higher level in the model, phonology (or phonemics), he demonstrates how the multitude of phones in the language are grouped by native speakers into abstract sets--*phonemes* (/ /). Using pairs of words such as *kill-gill*, *kill-hill*, in which each member of a pair has a different meaning cued by the minimal difference in the initial consonant, he shows that the initial consonants are sense-discriminative constituents, or phonemes. By such a technique, he groups phones into phoneme families; the members of the family are then called *allophones*. In (1), three voiceless back velar allophones of English are categorized under the /k/ phoneme. The description (Descrip) and distribution (Distrib) of each allophone are part of the analysis.

(1) /k/ voiceless back-velar plosive

[k<sup>h</sup>] Descrip: Voiceless aspirated back-velar plosive  
 Distrib: Obligatory before a stressed vowel, but is not used after [s]: *kill*

[k<sup>-</sup>] Descrip: Voiceless unreleased back-velar plosive  
 Distrib: Optional at the end of a word before silence: *sick*  
 Obligatory before another plosive stop phone: *active*

[k] Descrip: Voiceless unaspirated back-velar plosive  
 Distrib: Optional at the end of a word before silence: *sick*  
 Obligatory in all other environments  
 not mentioned: *skill*, *locket*

Taxonomic phonology also includes an analysis of the constraints on phoneme combination--*phonotactics*. For example, word-initial, two-member clusters with /k/ are limited in this way: /k/ may be preceded



only by /s/; /k/ may be followed only by /l,r,w,y/. Among three-member clusters, only /skr,skw,sky/ are permissible.

In the taxonomic model, the sound system is not completely described in the phonology. To get a fuller picture, although still incomplete, it is necessary to take another step up the structural hierarchy to the morphology.

At the level of morphology, the analyst studies phonemes and phoneme combinations which carry meaning, *morphs*. Morphs carrying the same meaning are grouped into families, *morphemes* ({ }) - the family members are called *allomorphs*. Sometimes, the distribution of allomorphs in words is conditioned simply by the words they are found in, as in the case of the {LEV-} allomorphs in (2)a. At other times, the allomorphs are conditioned by immediately contiguous phonological factors, as shown for the three {-Z} plural allomorphs in (2)b.

(2)a {LEV-} 'raise'

/lev-/ Distrib: *levy, levee,*

*levity, levitate.*

Optional in *lever,*

*leverage*

/liyv-/ Distrib: *alleviate.*

Optional in *lever,*

*leverage*

/ləv-/ Distrib: *elevate*

(2)b {-Z} 'plural'

/-s/ Distrib: after voiceless nonsibilants:

*sacks, maps*

/-z/ Distrib: after voiced nonsibilants:

*tales, eyes*

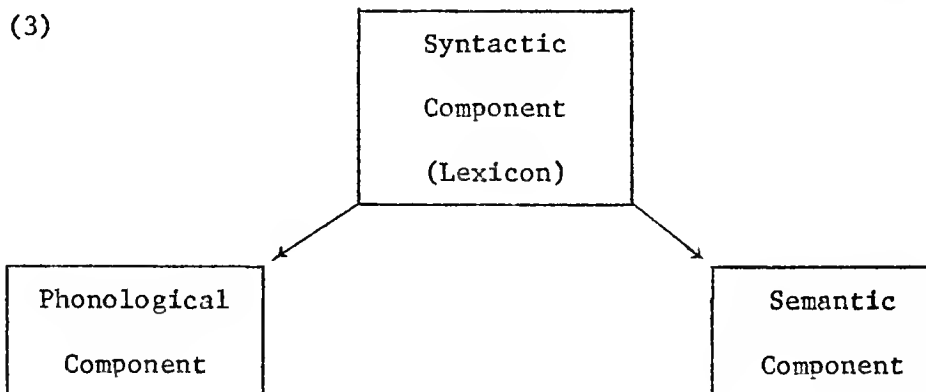
/-əz/ Distrib: after sibilants: *glasses, bridges*

As the examples in (2)a and (2)b show, phonemes of allomorphs take on different shapes when combined with other allomorphs. The study of such effects is called *morphophonemics*. This is a weak area of taxonomic research; except for cases of phonological conditioning, analysts in this tradition have not identified the language-wide principles which reveal how the different phonemes in a set of allomorphs, as in (2)a, are systematically related. Phonemes appear to be islands in words, their presence largely unpredictable. This research deficiency has earned for the taxonomic phoneme the name *autonomous phoneme*.

It is at the level of morphology that stress phonemes (´, ^, ˘, ˙) combine to form *superfix* morphs (e.g., ´ ˘, ^ ´), and pitch phonemes (1, 2, 3, 4) combine to form *intonation* morphs (e.g., 231, 232). The superfixes at the word level are considered unpredictable. But superfixes above the word level and all intonation patterns are thought to be systematic. However, the distribution of the superfixes and intonation patterns cannot be described at the morphemic level. Uncharacteristically, still higher levels of analysis are required, where syntactic notions like 'compound noun' and 'information question' are relevant.

In summary, the taxonomic theory says that the language is organized and can be described as a hierarchy of discrete levels built one upon the other from a phonetic foundation. As we have seen, it takes every one of these levels, from phonetics to syntax, to understand the sound system of English in taxonomic terms.

In the standard generative model, there are three parts to the grammar, a central syntactic component and two satellites, the phonological component and the semantic component, as diagrammed in (3). The sound system of a language is described completely in the phonological component.



A generative description of the sound system requires special data: The phonological component accepts as input a structurally analyzed string. As output it provides the 'phonetic representation' of this string. (Chomsky and Halle 1968:164)

Arriving from the syntactic component, the string consists of words, each labeled by part of speech and dissected according to its internal constituents--prefix, stem, suffix and other formatives. Each formative in a word is presented to the phonological component as a sequence of abstract phonemes.<sup>1</sup>

Phonological rules operate upon each word, transforming its abstract shape into a concrete, pronounceable shape. The sample derivation given in (4) shows how the rules systematically generate the phonetic forms of *elevate* and *levity*. The derivation begins with the phonemic representation of each word as input to the ordered rules. At intermediate stages in the derivation, each applicable rule alters the input. When all relevant rules have applied, the output is the phonetic representation of each word.

(4) /e+lev+at/	verb	← PHONEMIC INPUT →	/lev+iti/	noun		
é	ě	ā	← Stress Assignment →	é	ĩ	ĩ
			Final IOU Tensing →		ĩ	ĩ
		āy	← Diphthongization →		ĩ	ĩy
		ēy	← Vowel Shift			
		ě	← Vowel Reduction →		ě	
[é	lěv	ət]	← PHONETIC OUTPUT →	[lěv	ətĩy]	

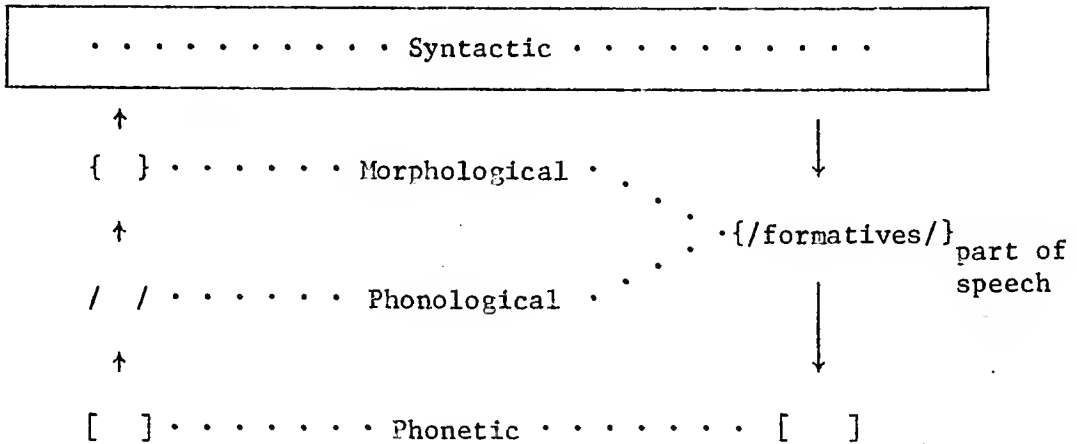
Note that the common semantic element in these two words is phonemically identical at the input stage, /lev/. At this level, semantic relatedness is mirrored in phonological relatedness. At the output stage, the transformational rules have produced the phonetically different [lěv] and [lěv]. A phoneme which can be converted systematically into different phonetic realizations thereby showing the nonrandom relationship between words is called a *systematic phoneme*.

In summary, the generative theory claims that the levels of language are interdependent. In the phonological component, where the sound system is described, transformational rules give meaning-bearing elements a sound shape; they capture the phonological relatedness of semantically related words, and they provide the means by which sound

can reflect information from all parts of the grammar. The transformational rules also assign stress and intonation to levels above the word, that is, to the whole input string.

A comparison of models is given in (5).

(5) Taxonomic Model                      Levels of Analysis                      Generative Model



The elements needed by each model to describe the sound system of a language are listed below.

Taxonomic Model

Generative Model

Syntax

Morphology

Intonation Morphemes

Superfix Morphemes

Morphophonemics for

{-Z} and {-D} Morphemes

Phonology

Phonotactics

Phonemes

Allophones

Phonemic representation of

a structurally analyzed

string of formatives

Transformational Rules

Phonetic Representation

The heart of the taxonomic sound system is the autonomous phoneme-- built up from phonetics and joined together to form larger units at higher levels. In the generative model, systematic phonemes in words

feed transformational rules which produce the correct intonation and superfix patterns, all consonant and vowel alternations and the appropriate phonetic forms to be spoken.

### IMPLICATIONS AND APPLICATIONS

Taxonomic phonology and generative phonology are not theories about language education, per se. But the implications and applications for pedagogical practice originating from these theories have profoundly affected ESL pronunciation courses. And not surprisingly, they have pointed educators in radically different directions. The concern of this section is to identify exactly what interpreters believe these opposing theories have to say about pronunciation teaching.

A theory of language, such as the ones outlined above, makes claims about language structure and about language users. The claims they make are relevant to language education because they say something about the linguistic content to be learned and about the learner's linguistic capabilities. Our specific topic then is this: How have the taxonomic and generative theories shaped pedagogical thinking about the linguistic content of a pronunciation course and about the learners in that course? The answer to this question will lead us, at the end of this section, to consider the goals--behavioral objectives--that we may reasonably expect the learner to achieve given his capabilities to handle a certain kind of linguistic content.

#### Linguistic Content

Consider the typical taxonomic-based pronunciation textbook. Its phonological content is shaped by two basic claims which the technical analysis makes about phonology. First, a phonology consists of a set of phonemes, each an abstract label for a family of phonetically similar phones described in terms of their distribution in the language. Second, the phonemic categories can be defined entirely without reference to information from any higher levels of linguistic analysis. That is, they are autonomous. Each of these claims has far-reaching ramifications for pedagogical materials.

Implication one. The claim that phonology consists of a set of phonemes implies that the principal content of phonological instruction is an inventory of those phonemes. However, the inventory of phonemes typically presented to ESL learners is not complete. In addition to vowel and consonant phonemes, the system also includes stress, pitch and juncture phonemes. However, juncture is considered a refinement too subtle to be relevant in pronunciation class. And stress and pitch phonemes are not taught as such. So, of the phonemes in English, ordinarily only the segmental (vowel and consonant) phonemes are presented to learners.

The phonology is not a rule system; it is a list of phonemes. Therefore, the phonological instruction which the learner receives is the following repeated point: This phoneme is one of the significant units in the sound system. The learner's attention is drawn to this fact by a symbol, usually from a linguistic symbol set.

The learner's phonological instruction is considered important: The inventory of phonemes identifies the full set of contrasts which are essential to make orally and aurally for English; it defines his phonological learning task. His phonetic instruction is also regarded as important. When he has learned to pronounce the common allophones of each phoneme in the inventory, he has the articulatory ability to pronounce any word in English. This skill is useful for the words he already knows, and for all words he does not yet know. For each of the latter words, however, he must first determine which sounds are needed. To do this, he will probably follow the usual suggestion made by taxonomically-oriented teachers, as stated by MacCarthy (1978:66): 'A pronouncing dictionary should always be at hand for consulting on which sounds occur in a given word.' Recognizing that a great deal of dictionary work lies before the learner, some pronunciation courses teach a symbol system adopted by a recommended dictionary.

Beyond phonology per se, a small amount of morphophonemics enters into the content of some pronunciation texts. The predictable allomorphic variants of the {-Z} and {-D} morphemes are presented together with their distributions. Students are asked to judge the voicing

characteristics of the final stem sound of a word in order to select the appropriate allomorph. Before the student can make such a judgment, however, he must have a phonetic description of each sound studied. This content is introduced in the phonetics instruction of courses in which {-Z} and {-D} allomorphs are taught.

In a word, the focus of pronunciation instruction influenced by the taxonomic model of phonology is ultimately on the *sound*, the phonetic realization of a phoneme. By oral-aural techniques, the learner is led to discriminate each major sound category from all others and to articulate each in a clearly recognizable way.

Implication two. The second major claim of the taxonomic theory is that phonemic analysis involves a studied nondependence on any higher levels of analysis. The pedagogical implication derived from this claim is this: Phonological instruction can (should) be conducted without reference to syntactic, morphological or lexical information. The effect of this implication has been profound: The autonomous phoneme has isolated pronunciation instruction from all other ESL teaching.

The theoretical stand of phonemic autonomy pervades taxonomic pronunciation texts. For example, since the phoneme is not a meaning carrier (as the morpheme is), the phoneme need not be practiced in a meaningful context. Although certainly in the minority, some writers have used nonsense words and even nonsensical sentences as exercise materials. Carrying this point even further is the general admonition to the teacher not to be concerned if students do not understand what they are saying and not to spend much time giving definitions or interpretations. And since the phoneme is defined entirely without respect to syntax, the phoneme need not be practiced in a grammatical context. In fact, word lists are the most common type of phonological drill material seen in older pronunciation texts. Contextualized pronunciation practice did not appear until the suprasegmental *morphemes* (superfixes and intonation patterns) were introduced. More recent texts have not held so rigidly to the word level for vowel and consonant work.

Phonological and phonetic instruction also carries out a strictly hands-off policy toward the spelling system, considered a derivative phenomenon of marginal interest. MacCarthy (1978:63-64) states this policy and one of its rationales:

It is a commonplace that the conventional spelling of English is inconsistent and therefore unreliable as a guide to the proper sequence of sounds to be articulated in speaking... The spelling of written English and the pronunciation of sequences of English sounds are two different matters, so different that they are best tackled separately and not, in the first instance, one in terms of the other.

Buttressed by the assumption that speech is primary and the presumably relevant fact that language is always learned natively in its spoken form before it is written, taxonomic pronunciation work is conducted exclusively by oral-aural means with little or no reference to written material. Except in the newest textbooks, spelling information, if given at all, has been treated as supplementary and tangential to the main thrust of the lessons.

As the phonemic system has been isolated from other grammatical subsystems, so phonological instruction has been largely isolated from other instruction in the ESL curriculum. When this analogy is carried out thoroughly, it does a disservice to the learner. To communicate, the speaker must manage concurrently all levels of linguistic analysis; they must be well integrated. But in many texts, the learner's phonological instruction does not recognize any other levels in the grammar. Therefore the learner receives little help answering a question of major import: How does my knowledge of the major sound contrasts in English interrelate with my knowledge of other subsystems in the language?

In summary, the phonological content of a typical taxonomic pronunciation book is the full set of autonomous vowel and consonant phonemes in English presented without reference to their spellings and exercised largely at the word level.



Now, let us consider a pronunciation text based on the generative model. Its phonological content also conforms to two fundamental claims made by the technical model about phonology. First, a phonology is a set of rules which converts an abstract form of a word (the input to the rule) to a concrete, phonetic form (the output of the rule). Second, the model says that a phonology is part of a highly integrated system in which syntactic, morphological and lexical information is relevant in phonological rules. For each of these claims there are major implications for teaching pronunciation.

Implication one. The first implication of the theory is that phonological instruction should involve the learner in using *rules* which apply to a *word* in its *spelled* form in order to *generate a pronunciation*. Several elements in this statement deserve discussion.

To say that learners should learn *rules* is to say something new and something old. What is new is that the learner's cognitive equipment is given a larger role in learning pronunciation than ever before. What is old is that the learner has long been expected to master rules affecting syntax and morphology. Rule learning in pronunciation courses was not intentionally neglected; until generative research began, there were no phonological rules to learn and only a handful of distribution rules in morphology.

It is highly significant that rules apply to *words*. Not only is the word the context in which phonological processes operate, but the word is also a communicative unit, unlike the phoneme. To treat the word as a message-bearing element points our materials away from nonsense words and meaningless sentences and toward the communicative function of language.

No large conceptual leap is needed to see that *spelled words* are the applied counterpart of underlying phonological representations. The abstract form of words is not available to the learner as an input to his rules, but an accessible surrogate does exist--standard spelling, widely acclaimed as remarkably similar to the ideal phonological representation of words found in technical analyses (Chomsky and Halle 1963: 49, 69, 80, 184n, 221). For this reason, spelling is a valuable tool

for the literate learner.

The use of spelling for pronunciation purposes not only acknowledges the fact that the written language is English, too--as all literates already know, but also emphasizes the important point that there is a very close link between written and spoken forms of English. That link is so useful that generativists can find no reason to separate the spoken and written modes of English even at the earliest levels of instruction. Wardhaugh (1974:168) says,

We must work out a system whereby from the first we can present our students with the sounds of English and the spellings of English *concurrently*.

Although the focus is on spelling, the learner's purpose in becoming familiar with spelling patterns is not to write properly, but to speak properly. The skill he is developing is receptive, not productive. Ideally, in an integrated program, both emphases--speaking correctly and spelling correctly--would have their place. In a pronunciation class, the student's main interest is to determine from the written form which sounds to say. In a spelling class, his focus shifts to the problem of how spoken words are represented in written form.

The centrality of spelling in pronunciation work points clearly to oral reading as the way the learner practices his phonological skills. Of course, he is also practicing his phonetic skills at the same time. As he reads aloud, he is matching his phonetic production to his phonological predictions.

By using rules, the learner can *generate a pronunciation* from a written form. This is a startling claim with revolutionary implications.

First, the ability of the learner to predict sound from spelling gives the learner a new status, one of independence. He no longer needs to carry a pocket-sized dictionary or depend on other crutches in order to figure out the pronunciation of an unfamiliar word. From his phonological instruction, he can determine for himself where the stress of a word belongs, which vowels will be long, short and reduced,

and which consonants should be used. Then, from his phonetics instruction, he can pronounce the sequence of sounds as predicted. Thus, using his rules creatively, the learner can generate a valid pronunciation for a word he did not know beforehand. This is an enormously valuable skill for one whose oral vocabulary should be constantly expanding.

The predictive value of a rule applied to spelling can be illustrated with the problem of determining which form of the plural morpheme to use for a word. Taxonomic materials typically make unreasonable assumptions and demands at this point. First, it is assumed that the learner knows the pronunciation of the word to be pluralized, and in particular what the final stem sound is. Second, the learner is required to learn enough phonetics information to judge the character of the final stem sound and determine which allomorph is needed. With *spelling* as a guide, rather than sound, the learner's decision process is simplified. The decision he must make is whether to add a 'long form' (/əz/) or a 'short form' (/s/ or /z/) to the word. By knowing that the /əz/ pronunciation is required after the spellings in (6), the learner will know to use a short form with all others. (No emphasis is placed on the distinction between /z/ and /s/ because the presence of the signal rather than the voicing of the signal is regarded as the main auditory clue of the plural.)

(6)	-se	-ze	-ge	-sh	-x
	-ss	-zz	-ce	-ch	

Second, the ability of the learner to predict sound from spelling gives even the pronunciation course a new status. Pronunciation instruction cannot be considered of minor consequence in the ESL curriculum, because it can now play a major role in the learner's ongoing language development. Because of his instruction, he takes away from the classroom more than articulatory and discriminatory skills: he takes away a generative capacity to use for the rest of his English-speaking career. A familiar saying can be recast to capture this point:

Teach a man the sounds of a word and he will be able

to say that word. But teach a man to predict the sounds of a word, and he will be able to say any word.<sup>2</sup>

In short, the first pedagogical implication of generative phonology--that instruction must involve rules which apply to spelled words for purposes of predicting the phonetic constituents of words--dramatically wrenches pronunciation teaching out of its customary role. First, it gives the class a *new orientation*. The focus is now on the *word*, not on the sound. Second, it puts a *new goal* before the learner. The learner's goal is no longer only the discrimination and articulation of sounds, it also includes the *prediction* of sounds in the context of words. Third, it engages the learner in a *new medium*. Now, in addition to oral-aural work, the learner does *visual-graphic* work as he makes use of spelling for pronunciation purposes. And fourth, it introduces *new content* into the course. Added to the usual concern for vowels and consonants, there is a new emphasis on predicting the *stress of words*, traditionally believed to be unpredictable. So, in orientation, goals, medium of learning, and course content, the concept of pronunciation instruction has been radically altered by the influence of the generative model.

Implication two. The second implication of the theory is that the isolation of pronunciation instruction within the ESL curriculum must end; the interrelatedness of linguistic subsystems is a fact which the ESL curriculum should reflect in its instruction for the benefit of the student.

The learner's phonological rules accept words which are 'structurally specified'. That is, the rules require syntactic, morphological and lexical information. There is, however, no realistic way that all the information necessary for pronunciation can be built into a pronunciation course. Instead, the course must depend for its success on the contribution of the whole curriculum. The grammar class must familiarize students with parts of speech and help them learn syntactic patterns. Early reading classes must acquaint students with the use of the alphabet and the reading process. Because of such help, pronunciation instruction can reciprocate: Oral work in all other

classes will be more effective; vocabulary building will be strengthened; reading skills will improve; and the appreciation of literature and poetry will be heightened.

A curriculum in which all the elements develop in a mutually supporting fashion helps the learner in a crucial way. The learner who has assimilated the language as a cohesive system of interrelated subsystems is better able to communicate, an activity in which multiple systems operate simultaneously to convey a message. In short, a well-integrated linguistic competence is prerequisite to a well-integrated linguistic performance. And the learning process can and should foster such an integration for the student at both the abstract and concrete levels.

In the pronunciation class, the integration is implemented by providing, in instructional materials, genuine grammatical contexts from which the students can gain the higher-order information they need for their rules. Word-level exercises have their place, especially in some phonetics instruction, but in phonological instruction, phrases, sentences and larger communication units are needed.

If the students are not getting the necessary support from elsewhere in the curriculum, the pronunciation course may have to supply the students with the required background. Ordinarily, students gain the ability to identify the syntactic categories of noun, verb, adjective and adverb in their grammar classes. But an ability to recognize morphological information such as prefixes and suffixes may not be nurtured anywhere in an ESL curriculum. This may then be an area where supplementary materials are needed in the pronunciation class.

In summary, the phonological content of a pronunciation text influenced by the generative model is basically a set of rules by which the learner predicts the pronunciation of words using all relevant linguistic information.

As this section has shown, the linguistic content of a pronunciation course is heavily shaped by its theoretical model, whether taxonomic or generative. Both models insist on articulatory phonetics for the concrete aspect of the course. But for the abstract aspect, the

two phonological theories carry learners in opposite directions. The point by point comparison of these directions makes clear that the generative theory provides the more substantial foundation upon which to build a pronunciation course.

### Learner Capabilities

The content of a pronunciation course is determined not only by what is known about the subject matter, but also by what is known about the learner. So, before asking what we will expect a learner to master from the course, let us consider what the learner can do. What are his capabilities?

The taxonomic and generative theories endow the language user (and by implication the language learner) with certain linguistic abilities. A consideration of these abilities not only accentuates the distance between the two theories but further explains their effect on teaching practice. It will also help to define more clearly the kinds of learning objectives we may expect the student to meet.

Judging from his model, we must assume that the taxonomist considers the learner's capability to include a well-developed memory, a linear assembling device and a rudimentary phone-matching facility. At the level of phonology, the learner uses his memory and his matching device; at the level of morphology, his memory, his assembling device and his 'matcher' come into play.

Since the constituents of a word are autonomous phonemes, they are by definition unpredictable. Anything which is unpredictable must be memorized. In the area of phonology, this means that the inventory of phonemes must be learned by memory. But in speech, the phonemes are represented by allophones. So, to select the appropriate allophone of each phoneme, the learner needs his phone-matching device, e.g., nasal vowels must be juxtaposed to nasal consonants, unaspirated voiceless plosives follow [s], etc.

In the area of morphology, the unpredictability of phonemes means that for any given word, its particular phonemes--the vowels, consonants and stress elements--and their proper order must be memorized. Then,

when the learner uses English words, his linear assembling device puts together the constituents of a word according to his memory of them or according to the guidance of a dictionary. Occasionally, he may have to use his matching device to pick the correct allomorph to add to a stem.

The implication of the above is that the learner's phonological task in pronunciation class is to learn the phonemes. And his phonetic task is to learn the allophones of each phoneme and where to use them. Strangely, here is one of the few places where an implication of taxonomic theory has not been carried out with meticulous care. On the basis that differences among allophones of the same phoneme never distinguish two words with different meanings, allophones are generally considered irrelevant to communication and therefore of no importance to the learner. Thus, except in the occasional text, the learner is given no explicit guidance for the use of allophones. And his oral practice is rarely designed so as to group similar allophones together. Instead, as he concentrates on a phoneme in his drillwork, he is barraged with the full range of its allophones without regard to their difficulty. Presumably, if the learner can articulate any one of these allophones, he has satisfied the teacher's expectations. This approach to phonetics instruction, where all attention is centered on the abstract phoneme and none on the concrete allophones, implies falsely [1] that there are no allophonic differences, an implication which confuses the learner whose non-English ear hears the differences, and [2] that the word contributes nothing to the spoken shape of a phone. Indeed, when words are used in instruction, they are merely the incidental framework for phonetic practice on target phonemes.

The learner commits most words to memory as wholes. He learns each word as a sequence of autonomous phonemes, whether it consists of a single morpheme or of multiple morphemes. Derivational affixes are considered permanent constituents of a word. Their morphophonemics play no role. That is, the learner makes no choice among *con-*, *com-*, and *col-* on the basis of environment for words like *convene*, *composer* and *collection*.

Inflectional affixes are another matter, because they are not invariably present on a word. For them, morphophonemic considerations enter into the learner's concatenation process. Of the eight inflectional morphemes, only the three {-Z} morphemes (plural, genitive and third person singular present tense) and the two {-D} morphemes (past tense and past participle) have alternate forms which are phonologically conditioned. For these morphemes, the learner must use his matching device to select an appropriate allomorph.

When the learner puts autonomous phonemes together to form words such as those in (7), he is operating at the morphemic level. At this level, of course, the allomorphs of the morpheme {MIN-} are related by semantic criteria. However, even though the allomorphs (/mayn-/ , /min-/ and /mən-/) look much alike in spelling, their phonological differences are considered unrelated, that is, entirely due to historical accident. The learner sees no special interconnections among the autonomous phonemes, /ay, i, ə/.

(7) /ay/	/i/	/ə ~ ay/
mínör	mínimál	minóřitř
mínüs	míniscùle	minútiã

This fact has its pedagogical ramifications for the learner and the teacher. Since the varying phonological shapes in the word set are independent of each other, the learner must memorize the phonemic constituents of each word as if the words were not related at all. There is no phonological advantage in learning the words in sets. Furthermore, since the learner sees no phonological relationship among the varying phonemes in a morphemic set, there is no motivation to present the phonemes to the learner in sets or in any particular order. The teacher has complete flexibility in sequencing the presentation of phonemes.

In short, taxonomic phonology says to the learner: 'Your memory, your assembling device and your matching facility are adequate for memorizing the phonemes of English and for combining them into words.



No greater facility is necessary because no greater structure is present.'

In the generativist's view, the learner has an impressive store of cognitive equipment to use for learning another language, or even, the phonology of that language. What sort of abilities does he have?

To be sure, the learner is capable of learning the major contrasts which must be signaled in the language through sound. But the signals bear so much information that no ordinary memory or adding machine could manage the load. The learner's equipment must handle rules which organize the details of each subsystem and relate them to other subsystems. In phonology, this means that the learner's equipment is adequate to use sounds in systematic ways [1] to signal their place in the structure of a word, [2] to highlight the sound relationships obtaining between different words, and [3] to convey (often redundant) information from other levels of structure. (And all this is done while simultaneously generating an overarching rhythm of alternating vowel qualities and stress levels characteristic of a stress-timed language [Dickerson 1978].) For this kind of cognitive processing, only the power of a sophisticated, rule-using computer could serve as a model.

It is tempting to think, as taxonomists do, that phonemes are added together to make up a word and that a sequence of appropriate allophones gives a word its sound shape. But generativists claim the opposite, namely, that in reality words give sounds their shape. That is, sounds take their character from their special place in a word--in relation to the vowels, consonants, prefixes, endings, stresses of a word. For example, to the left of adjectival *-ic* the stressed V in a VC sequence will be lax, e.g., *acrobatic*; the only kinds of sibilants permitted before the *-ion* ending are the palatals, /š. ž. č. j/; degrees of stress have predictable positions in a word depending on the prefixes, suffixes, vowels and consonants in a word; and only long vowels may appear immediately before a word-final /ð/, e.g., *bathe*. Thus, words put limits on the kinds of sounds which can occupy different positions in the word. Sounds so limited in these ways therefore signal to the ear a message about their relationship to neighboring sounds and

to the word as a whole (Dickerson 1980).

Orderly, word-level constraints on sounds are, of course, the rules by which the learner can predict those sounds. Therefore, to teach predictive skills, practice materials must focus on written words, and words which have enough structure so that the rules can be exercised.

With sounds, the learner can signal intra-word relationships, as just shown. But he can also signal inter-word relationships. His ability to use rules allows him to know, for example, that the /ay, i, ə/ constituents in the {MIN-} derivatives above do not form a random set. In fact, the trio says to the ear of the listener, 'we are related in a special way.' Similarly, the learner's rules tell him that the [k, s, ʃ] of *electric, electricity, electrician* are systematically linked to each other and, with his rules, he can generate the appropriate phone for each word, although all are spelled with the letter *e*. Thus, the learner can use rules to signal the relationships which listeners expect to hear within a word and between words.

The fact that the learner is capable of seeing the inter-word relationships of sounds in each set has a bearing on teaching materials. First, word sets are highly useful in exercises. The words in (8) illustrate again that the particular phonetic shape of each set member arises from the special place each vowel and consonant holds in its own word. When the student learns to generate the proper variant by rule, he should be given such word sets as practice material. Without a knowledge of the necessary rules, there is little point in trying to teach the variants by simply presenting the learner with lists of related words.

(8)	/ay, i, ə/		/k, s, ʃ/
	final	derive	politic      mercantile
	finish	derivative	politicize      commerce
	infinite	derivation	politician      commercial

The advantages of using word sets with learners who know the rules

are numerous. First, such practice builds the learner's confidence in the orderliness of the sound system and in his ability to use that orderliness to predict the sounds of words in derivational sets. Second, prediction of dissimilar phonetic shapes from similar spelled shapes helps break down a common learner misconception about English words, namely, 'if they are written alike, they must sound alike'. Third, for each spelling, such as *i* or *e*, there is a limited but highly stable range of regular sound correspondences. Practice using word sets promotes a familiarity with this range and helps move the learner from mechanical prediction to a more automatic response to letters in certain environments. Fourth, practice with word families helps the learner expand his vocabulary in an ideal way. He learns new words in relation to old words, and at the same time learns how to pronounce the new words by rule instead of by memory. In effect, we are giving the learner practice, in the classroom, doing the kind of analysis we hope he will do when he leaves our classroom.

The second pedagogical implication of the fact that phonetic relationships exist among the allomorphs of a morpheme is this: The presentation of phonemes in a class should anticipate the use of predictive rules. For example, before teaching the effect on consonants of a final *-ian* ending (e.g., *physician*, *technician*, *optician*, *phonetician*), the teacher should have done preparatory phonetics work not only on the /ʃ/ phoneme, but also on the /k/ phoneme. In this way, practice with the consonant rule using word sets will be more effective, since the /ʃ/ words are used in conjunction with related /k/ words (e.g., *physical*, *technical*, *optical*, *phonetics*). Thus, the predictability of word relatedness should affect the order in which phonemes are introduced for phonetic practice.

One final relationship between phonology and the larger system bears some mention. So that the message is not easily lost, language is full of redundancy. The phonology, in particular, is full of echoes of morphological, syntactic and lexical information. For example, to know whether to render the ending of the word *duplicate* as /ət/ or as /eyt/, the learner must know whether or not the word is a verb. His

phonological rule says that a verb requires /eyt/, while the noun and adjective require /ət/. By using the syntactic clues surrounding the word, the learner can determine its part of speech and make his decision. When the word is in context, the proper ending simply reinforces the syntactic evidence already available; it is redundant. But when the word is out of context, or its context is not clear, the phonetic signal may stand alone to interpret the syntactic category of the word: [dúwplěkěyt] is a verb.

Or consider the orthographically ambiguous sentence in (9). When this sentence is spoken, the phonological rule disambiguates the verb, because the prefix meaning 'again', which attaches only to free stems, must carry a major stress and a full, tense vowel. The lexical information in the phonological rule is reflected back in the phonetic output. In this sentence, the lexical information in the phonology is critical; in a larger context, however, it may be redundant.

(9) He restored the furniture.

[rřstórd] = 'renovated'

[ríystórd] = 'stored again'

In the generative system, phonological rules reveal many relationships: sound to sound in a word, sound to sound between words; sound to morphological, syntactic and lexical levels. Sounds carry many messages from many levels of the system. The learner is capable of using this system of interwoven subsystems by learning the interpretive rules and using them on spelling. The results are predictions which appropriately carry the expected messages.

The rules which the learner knows and uses to interpret spelling are not simply superficial, spelling-to-sound correspondences, as in the phonics rhyme for long vowels: When two vowels go walking, the first does the talking. Some of the rules are also *language* rules apart from spelling. That is, the learner's rules concerning vowels, consonants and stress capture significant generalizations about English (many of which were unknown prior to generative research). Since

spelling represents words largely at the systematic phonemic level, the pedagogical rules will generate for the learner the same sound-system patterns that native speakers use. Thus, the learner's rule practice does not merely establish spelling-sound links; it also accustoms the learner to the structure of sound at the word level--precisely what the phonology says the student should be learning.

Assuming, as discussed at the outset, that the purpose of phonological instruction in the pronunciation class is to foster a native-like competence in the phonology, we have undertaken to examine the native competence and to extend its capabilities to the language learner. What can the learner do, and what kind of demands can be made of him? The taxonomist assumes that the learner has only minimal cognitive equipment--on the order of an adding machine with a memory--and therefore asks him to learn little. The generativist assumes that the learner has a tremendous cognitive capacity--like that of a great computer--and therefore asks him to exercise it to the full.

In the last two sections, the rule system characterizing English phonology has been shown to be complex and demanding. To handle such a system, the learner must have more than the kind of competence apparently required by the taxonomic model. To be adequate to the task of acquiring a native-like control of the English sound system, the learner must have the potential of developing a sophisticated competence like that attributed to language users by the generative model.

But how much of the ideal competence can we reasonably expect the learner to acquire? This question is broached in the next section.

### Behavioral Objectives

In the early 1950's, language educators turned to phonologists for guidance with pronunciation instruction. The taxonomic model adopted at that time for ESL teaching is still dominant in pronunciation textbooks. This model, however, not only embodies an inadequate view of phonology and an impoverished concept of the learner, but in many respects it has also led educators in unfortunate directions. Now, in the 1980's, we must turn again to phonologists for help. As shown in

the sections above, the current generative model provides a more satisfying description of the sound system and attributes to native speakers a more realistic linguistic capacity. When implemented in the classroom, this model leads the learner to a more complete attainment of the pronunciation goals set out at the beginning of this paper:

- to gain concrete performance skills in the area of phonetics, and
- to gain an abstract native-like competence in the area of phonology.

The generative model is, therefore, a much more adequate basis for a pronunciation course than the taxonomic model. For this reason, the behavioral objectives outlined below reflect the generative orientation to phonology and the generative assessment of native-speaker competence.

Any pronunciation course modeled on a linguistic description of the sound system will have a dual focus--phonetics and phonology. Furthermore, regardless of orientation, such a course will include the objectives under I. below, where discrimination and articulation of phonetic detail are dealt with. But if the course incorporates the generative model, its phonological component will concentrate heavily on the use of rules to predict the segmental and suprasegmental elements of the sound system.

By the end of a pronunciation course informed by generative phonology, the learner should be able to do the following things:

#### I. Phonetics - Discrimination and Articulation

##### A. Segmentals

1. discriminate all vowel contrasts
2. articulate all vowel segments
3. produce a pattern of alternating full and reduced vowels in words and phrases
4. discriminate all consonant contrasts
5. articulate all consonant segments

##### B. Suprasegmentals

1. discriminate stressed from unstressed syllables
2. produce a major stress in words, constructions, utterances

3. discriminate intonation contours

4. produce intonation contours

## II. Phonology - Prediction

### A. Segmentals

1. predict from spelling all predictable stressed and unstressed vowel sounds in words

2. predict from spelling all predictable consonant sounds in words

3. predict from spelling the predictable allomorphs of the {-Z} and {-D} morphemes for any given word

### B. Suprasegmentals

1. predict from spelling and other structural information the placement of predictable word stress

2. predict from syntax the placement of construction and utterance stress

3. predict from syntax and context the use of intonation contours

4. predict from syntax and context the use of contrastive stress

The objectives above are stated in general terms. The specific content matching some of these objectives is common knowledge. For example, the content based on the phonetics objectives can be found in any good pronunciation textbook (e.g., Dickerson and Dickerson, forthcoming). For other objectives, the specific content is not well known, because it involves recent interpretations of generative phonology. The interpretations will be identified in subsequent papers of this series. In particular, the papers will discuss the pedagogical rules by which word stress, and vowel and consonant segments can be predicted from spelling.

In broad terms, we have now answered one of the central questions posed at the outset of this paper; What can be done to incorporate new linguistic developments into our pronunciation course? The answer suggested here involves a three-part approach. First, we must recognize what the technical insights are and what they imply for teaching. That is, what does linguistic research say about the language system

and about the language user that is relevant to teaching? Second, we must modify the behavioral objectives of the pronunciation course to incorporate the values which can be gained from current phonological research. Third, we must translate technical insights into pedagogically useful materials in order to implement the new objectives. In this paper, we have dealt with the first two parts of this approach. In addition, we have pointed to subsequent papers in this series for a detailed treatment of the third part.

We close this discussion with a look at the second question raised at the beginning, Why has pronunciation teaching been so slow to keep pace with the growth of phonological information?

#### PAST AND FUTURE

At this writing, the taxonomic model of phonology governs the great majority of ESL pronunciation teaching. However, the categorical adoption of this thirty-year-old model is now past. The generative model of phonology, now fifteen years old, is beginning to have an effect on language pedagogy. This section concerns [1] the reluctance of ESL educators to recognize the inadequacies of the taxonomic model and the potential value of the generative model, and [2] some of the hopeful signs of a future for generative phonology in TESL.

To help with this discussion, we will refer to the results of an informal survey conducted on the tenth publication anniversary of Chomsky and Halle's *The Sound Pattern of English* (1968). Surveyed were teacher trainers in graduate TESL programs in the United States. Of the sixty forms distributed, I received responses to nearly half. As stated in the questionnaire, the purpose of the study was 'to survey the impact that generative phonological research has had on ESL teaching and teacher training.' The findings of this study are incorporated into the following discussion.

There are two principal reasons that generative phonological research has not yet greatly affected ESL pronunciation instruction. First, the taxonomic model defined pronunciation content for ESL: basically the segmental phonemes of English. This definition, with its



linguistic stamp of approval, set a precedent which has remained virtually unchallenged for three decades. The power of this precedent is demonstrated by the fact that during this time, all published pronunciation textbooks have followed the essentials of the taxonomic model. These textbooks have determined the course content for untrained teachers. But even among trained teachers, course content has been unchanged; most teacher trainers know only the taxonomic model. Therefore, across the field, this model is considered adequate.

Second, the generative model, in its complex, technical form, looks forbidding and impractical. Distinctive feature notation, the mathematical symbology of rules, and abstract underlying representations have no immediate or transparent applications. Nor is it clear to many where such applications might fit into teaching. Do they replace or do they supplement what has been done? A comprehensive, integrating theory of pronunciation resolving the matter is not well known. Therefore, in the considered opinion of some ESL researchers, generative phonology has little or no value in the classroom (Hammerly 1973:439; Wilkins 1972:67; Brown 1975:113). Furthermore, no pronunciation materials based on the generative model have been published to dispute this assessment. Thus, in the eyes of many, the new model offers no convincing alternative to the old model.

Despite this generally negative view regarding generative phonology in ESL, there are some positive indications that a change is coming. A few far-sighted individuals in the field have been able to envision the potential that generative work holds for pronunciation teaching (Wardhaugh 1974). Going further, specific suggestions for how the model can help the learner have appeared in professional journals (C. Chomsky 1970; Dickerson [cited below]; Kreidler 1972a,b; Schane 1970; Schnitzer 1970). Moving further still, actual teaching materials are beginning to be written. Schnitzer (personal communication) has completed a programmed textbook for teaching word stress, vowel quality and consonant choice to foreign students. At UCLA, a set of materials, 'Course in Advanced Pronunciation', was developed on a trial basis to teach ESL students how to predict the sounds of multisyllabic words.

And our own materials (Dickerson, forthcoming) offer additional testimony to the applicability of generative insights.

This literature is meager, but in its own way, it is helping to break down some barriers to the acceptance of applied generative research. Dickerson (1980) shows that distinctive feature representations, mathematical conventions, and many structural complexities are unnecessary for formulating pedagogically useful rules. Furthermore, when standard orthography substitutes for underlying forms, the input to the learner's rules is no more abstract than orthography (Dickerson and Finney 1978). Generative phonology has an important place in pronunciation teaching. It does not conflict with or supplant traditional concerns for articulatory work. Instead, it provides a long-needed adjunct to that work because it shows the student where in words he can use the articulations he is mastering (Dickerson 1977). Thus, there is no habit formation vs. cognitive code debate involved. Developing articulatory habits is necessary in phonetics instruction, and learning predictive rules is necessary in phonological instruction. Each complements the other and both are integral parts of a comprehensive, generative-based theory of pronunciation teaching as formulated in this paper and in Dickerson (1978).

The majority of respondents in the survey felt that generative phonology is important, but at the moment it is of greater value for teacher training than for ESL classroom use. In fact, in nearly all the programs where a course in phonology is required of Masters candidates, some time is given to the generative model. Cited as benefits of such study were the following points. Teachers should be aware of contemporary approaches to phonology; they should understand how a sound system works as an aid to analyzing student problems; they should have enough background in generative theory to keep up with current phonological research; they will find the generative analysis helpful in seeing the regularity of English orthography. For whatever reasons teacher trainees are exposed to generative phonology, it is encouraging to find such exposure. Along these lines, Kreidler (personal communication) is preparing a text on generative stress and intonation to be

used by ESL teachers and teacher trainers.

In short, the profession is slowly becoming aware of generative phonology, as evidenced by courses required of teachers-in-training. Applications of generative insights are also becoming visible through articles in the professional literature. And soon, there will be actual teaching materials available to give substance to claims that generative phonology in pedagogical form has a vital role to play in the ESL curriculum. Thus, there is reason for some optimism about the future of generative phonology in TESL.

#### SUMMARY

The content and goals of pronunciation teaching ultimately rest on and conform to certain theoretical linguistic foundations. This paper has sketched two phonological theories--taxonomic and generative--upon which such teaching is based. Both theories provide the learner with phonetics and phonological guidance. But do both do so adequately?

To answer this question, the discussion first isolated the differences between the two theories with respect to the linguistic phenomena each can explain and the concept each has of the language user. Then, the discussion highlighted and evaluated the pedagogical consequences of each theory. On numerous counts, the analysis led to a rejection of the taxonomic model in favor of the generative model as the superior foundation for pronunciation teaching.

Therefore, when the objectives of a pronunciation class were stated, they were defined in such a way as to incorporate the special contribution of the generative perspective. That special contribution is an insistence on *prediction*, that is, on the use of rules to generate the segmental and suprasegmental elements of an utterance. The nature of these pedagogical rules is the subject of the various installments in this series.

The paper ends on the optimistic note that a quiet revolution in pronunciation teaching is indeed taking place--one that will prepare the learner for his task as a literate speaker of English more adequately than ever before.<sup>3</sup>

*FOOTNOTES*

- <sup>1</sup>Each phoneme is represented as a set of distinctive features, such as [-syllabic, +strident, +anterior, +coronal, -voiced] for /s/. In our discussion, however, we will use unitary symbols, e.g., /s/, instead of feature sets to refer to abstract phonemes.
- <sup>2</sup>Give a man a fish and he will eat for a day; teach a man to fish and he will eat for a lifetime.
- <sup>3</sup>I wish to thank Lonna J. Dickerson and Rebecca H. Finney for the helpful criticisms and suggestions they offered during the preparation of this paper.

## REFERENCES

- Brown, Gillian. 1975. Phonological Theory and Language Teaching. *Papers in Applied Linguistics*, ed. by J. P. B. Allen and S. Pit Corder. London: Oxford University Press, 98-121.
- Chomsky, Carol. 1970. Reading, Writing, and Phonology. *Harvard Educational Review* 40:287-309.
- Chomsky, Noam and Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper and Row.
- Dickerson, Lonna J. and Wayne B. Dickerson. forthcoming. *Learning English Pronunciation, Volume I, Vowel and Consonant Sounds and Spellings*.
- Dickerson, Wayne B. 1977. Generative Theory in TESL Practice. *The Modern Language Journal* 61:179-187.
- Dickerson, Wayne B. 1978. English Orthography, A Guide to Word Stress and Vowel Quality. *International Review of Applied Linguistics* 16:127-147.
- Dickerson, Wayne B. 1980. Bisyllabic Laxing Rule: Vowel Prediction in Linguistics and Language Learning.
- Dickerson, Wayne B. forthcoming. *Learning English Pronunciation, Volume II, Word Stress and Vowel Quality*.
- Dickerson, Wayne B. and Rebecca H. Finney. 1978. Spelling in TESL: Stress Cues to Vowel Quality. *TESOL Quarterly* 12:2:163-175.
- Francis, W. Nelson. 1958. *The Structure of American English*. New York: The Ronald Press.
- Hammerly, Hector. 1973. Teaching Pronunciation and Generative Phonology. *Foreign Language Annals* 6:487-489.
- Kreidler, Charles W. 1972a. English Orthography: A Generative Approach. *Studies in Honor of Albert H. Marquardt*, ed. by J. Alatis, Washington, D.C.: TESOL, 81-91.
- Kreidler, Charles W. 1972b. Teaching English Spelling and Pronunciation. *TESOL Quarterly* 6:1:3-12.
- MacCarthy, Peter. 1978. *The Teaching of Pronunciation*. Cambridge: Cambridge University Press.
- Schane, Sanford. 1970. Linguistics, Spelling and Pronunciation. *TESOL Quarterly* 4:137-141.
- Schnitzer, Marc L. 1970. Applied Generative Phonology: A Methodology for Teaching Pronunciation. *International Review of Applied Linguistics* 12:4:289-305.
- Wardhaugh, Ronald. 1974. *Topics in Applied Linguistics*. Rowley: Newbury House.
- Wilkins, D. A. 1972. *Linguistics in Language Teaching*. Cambridge: MIT Press.



## TOWARD DEVELOPING AN OPTIMAL TEST OF WRITING PROFICIENCY

Cynthia A. Gockley

This paper reports the results of a preliminary study that was designed and carried out to help construct an instrument for testing writing proficiency in the ESL classroom. The ability to write embraces such a wide variety of skills that the scorer finds weighting the various skills a severe obstacle. The goal of our study was to make this task more manageable. An attempt was made to create a test instrument that combined objective, discrete point elements with the more subjectively scored composition element. A substantial correlation between each element with the total test score would assure that the total test was truly integrated. Too high a correlation between any two objective elements would suggest that the test elements were redundant. Keeping these considerations in mind, a test was constructed which had four major elements: punctuation, sentence ordering, cloze and composition. The test was administered to ten students of the Intensive English Institute in the Spring of 1979. The results show that sentence ordering, cloze and composition are appropriate for testing writing proficiency.

### INTRODUCTION

The problem of evaluating second language competence has by no means been resolved satisfactorily yet. In the existing instruments for testing, writing proficiency has been particularly neglected. Yet it is upon this skill that much of our students' future academic and professional success depends. Asking students to write a guided composition (limited and guided in terms of subject matter, time constraints, etc.) has been the traditional means of evaluating all-over writing proficiency: its appeal lies in its undeniable face validity, and the minimal time in construction. Oller, among others, states a strong preference for integrative tests of this type, a preference based upon current linguistic evidence that they are "more effective devices for eliciting information concerning the efficiency of the learner's internalized grammar" (Oller 1976:25). Discrete-point tests are limited in their ability to elicit the "global" writing proficiency of students. The pragmatics of evaluating such necessarily subjective tests present a formidable problem, however. Essays,

themes, and compositions are notoriously unreliable and time-consuming to score.

The ability to write embraces such a wide variety of subskills that the scorer finds weighting the skills a severe obstacle: The writer must apply more than a knowledge of English structure at the sentence level. He must possess adequate vocabulary to discuss the material precisely and appropriately. He must be familiar with the sheer mechanics of written English: its graphemic system, spelling patterns, and the grammatical function of punctuation. Moreover, proficient writing demands a command of the rhetorical devices of English, including a grasp of register and style, and the ability to organize and edit material in presenting a cohesive argument or description.

Attempts to resolve the issue have resulted in the use or suggested use of "eclectic designs" which combine the features of what Lado terms the "synthetic approach for wider sampling and easier scoring" with the "more readily acceptable validity" of composition. Harris and Rivers both agree that the ideal method of measuring writing skills is with a "battery of tests in which objective assessment (predominates)" (Rivers 1968:305). Harris claims that "well-constructed objective tests...have been found to correlate quite highly with general writing ability, as determined by the rating of actual samples of free writing" (Harris 1963: 71). The problem, then, was to define and create just such a battery of tests.

The goal was to create a test which combined objective, discrete-point elements with the more subjectively-scored composition element. On the one hand, a substantial correlation between each element with the total test score would ensure that the total test was truly integrated. This eclectic approach, as Lado points out, "does not guarantee a good test, since the total test will only be as good as the sum of its parts, and any part that is unsatisfactory will weaken the total test accordingly" (Lado 1961:251). On the other hand, too high a correlation between any two objective elements would suggest that the test elements were redundant, and would render the total test inefficient.



## PROCEDURES

Subjects. The subjects of the test were ten students of the Intensive English Institute's Structure 252A component, Spring 1979. The students were a well-motivated group; all intended to pursue academic work at the university-level within the United States. They understood the need to sharply ungrade their writing skills; many were under time constraints because of limitations on government scholarships. This affective factor resulted in a high degree of cooperation concerning writing assignments and attendance throughout the semester. Consequently, the students had had a good deal of practice in the kinds of skills that were tested. Nothing in the testing method was new to them; they had performed similar tasks in and outside of class, and only the content and complexity were different.

The test was administered two weeks before the academic term ended, and the students were informed that the test was to be graded upon the same basis as any other exam given during the course; this was to provide motivation to perform to the best of their ability without subjecting the students to undue pressure or self-consciousness which could have caused "test anxiety" and rendered the results less reliable.

The subjects were given two full hours in which to complete the test. About half of the group completed the test within the first hour; several (weaker) students spent the full time allotted, and only then reluctantly submitted their papers. This was not a power test; students were given ample time in which to perform at their highest capacity.

Instruments. The test itself was composed of four major elements, designed to elicit a representative sample of skills, although some of these skill areas necessarily overlapped. (See Appendices 1-3 for copies of the objective test elements.) The first three elements were based on a discrete-point approach which allowed the scorer to mark these items objectively. The fourth element consisted of a short composition which covered those skills that, according to Harris, combine "a number of diverse elements, only some of which are strictly linguistic."

Punctuation Test. Punctuation must be recognized as an important writing subskill. As Quirk points out, "Punctuation practice is governed

primarily by grammatical considerations, and is related to grammatical distinctions" (Quirk 1972:1055). Students were asked to punctuate a passage adapted from a recent "Newsweek" article. They were required to provide capitalization, periods, commas, and quotation marks where required.

A certain degree of flexibility was allowed: The use or neglect of a comma may often be a reflection of personal style. What was of concern here was the application of essential punctuation, those visual devices which perform a clear syntactic role: for example, apostrophes to indicate genitive rather than plural endings.

There were 52 points in the passage where punctuation was essential. Students were penalized for excessive punctuation or ungrammatical punctuation marks, but not where punctuation was variable. Each punctuation mark was equally judged: each was worth .5 point, making a 25.5 total. (See Appendix I.)

Sentence Ordering. The second section concentrated on student sensitivity to reference and linkage, and the ability to organize units of information to comprise a logical and cohesive paragraph. This was measured by means of a sentence ordering task. Sentence ordering tasks such as this are still rather controversial. Many sources state that the validity of such exercises in testing organizational skills employed in writing is dubious. Heaton, however, feels that such tasks have their place at the more elementary levels, where students are presented with unequivocally linear presentations. I felt that it would be interesting to see how such a manipulatory task would correlate with more creative aspects of writing (e.g., the composition component).

Students were given three points for each sentence that followed the correct preceding sentence (even if the preceding statement was incorrectly ordered in relation to its preceding statement). With nine sentences there was, therefore, a total possible score of 27 points. The passage was adapted from an anthropology text (Leakey, Richard, Origins) the students had been introduced to readings from the same text in class, although this particular passage was new to them. Therefore, the vocabulary was familiar to them. (See Appendix II.)

Cloze. The third section was an open-ended cloze test. More than one answer was acceptable on a number of items (see Appendix III, Answer Key). The cloze has been demonstrated to correlate highly with a wide variety of language proficiency skills, including but by no means restricted to oral fluency and the ability to read aloud. It has been suggested that clozes measure students' internalized grammar. The cloze was constructed on a random-deletion basis, with roughly every fifth word omitted, except where it was clear that deletion would elicit too many possible alternatives, or otherwise create confusion (for example, deletion of a word essential to the meaning of the entire passage which could not be "discovered" in the context of the passage). There were 42 items, each worth one point. The passage was adapted from a British ESL text commonly used by the IEI Structure components as supplementary material (O'Neill, English in Situations); this passage had, however, not been used in the classroom. (See Appendix III.)

Composition. In the final section of the test, the students were asked to write a short composition. There were several questions at the top of the page to guide students in discussing the assigned topic. The students were instructed to write neatly, concisely, and descriptively, and to take care to form a tight, cohesive paragraph that would be interesting to an outside reader. The topic ("My Favorite Month") was purposively innocuous and considered by the test designer to be relatively free of cultural bias. It was a subject that had been indirectly and casually discussed in class on several occasions: the students were familiar with the vocabulary and idiomatic structures needed to write about the subject, yet it was not a topic that had been assigned as a composition before. The topic was apt to produce sentences in simple present; it was hoped that students would not attempt to be overly-ambitious syntactically. As it turned out, most were not.

The composition was scored by four different graders. Two were teachers of this particular group of students: two were teachers of other classes within the IEI. Although the compositions were numbered, and graders theoretically did not know the names, ages, sex, background or standing of the students, bias on the part of the two 252A teachers

was, perhaps, inevitable; both tended to rate the compositions consistently higher than the more objective teachers did.

The graders were asked to rate the compositions on the basis of 1) content (ideas, substance, whether the student fulfilled the assignment or not), 2) form (organization of material, 3) grammar (syntactic patterns: used correctly and where required), 4) style (appropriate register, precision of meaning and expression), and 5) mechanics (including superficial acceptability of the composition, such as handwriting, spelling, punctuation, and layout). These features were based on Harris' answer to the question, "What constitutes good writing?" (Harris 1969). Graders were asked to use their initial impressions, and not to dwell overly long on the papers. The scores were then averaged, with a possible total score resulting of 50.

## RESULTS

The results of the test battery were viewed in terms of their correlations with each other, as well as with the total test score. Reliabilities were also examined. Distributions and other statistics for each subtest can be found in Tables 1-5.

Table 1

### PUNCTUATION TEST

#### Score Statistics and Reliability

k = 52	Reliability after applying the
N = 10	Spearman-Brown Prophecy Formula:
Range: 8.5 - 22 (14.5)	k = 100
Median = 17.25	r <sub>tt</sub> = .68
Mode = 21.5	
$\bar{X}$ = 16.45	
S.D. = 4.8	
K.R. <sub>21</sub> = .522	
S.E.M. = 3.32	

Table 2

## SENTENCE ORDERING TEST

## Score Statistics and Reliability

k = 27	Reliability after applying the
N = 10	Spearman-Brown Prophecy Formula:
Range: 0 - 24 (25)	k = 100
Median = 15	$r_{tt} = .974$
Mode = 0	
$\bar{X} = 13.8$	
S.D. = 7.37	
K.R. <sub>21</sub> = .91	
S.E.M. = 2.21	

Table 3

## OPEN-ENDED CLOZE

## Score Statistics and Reliability

k = 42	Reliability after applying the
N = 10	Spearman-Brown Prophecy Formula:
Range: 16 - 35	k = 100
Median = 26	$r_{tt} = .869$
Mode = 0	
$\bar{X} = 25.8$	
S.D. = 5.42	
K.R. <sub>21</sub> = .676	
S.E.M. = 3.08	

Table 4

GUIDED COMPOSITION  
Score Statistics and Reliability

k = 50	Reliability after applying the Spearman-Brown Prophecy Formula:
N = 10	
Range: 28.5 - 45 (17.5)	k = 100
Median = 34.13	$r_{tt} = .73$
Mode = 32.5	
$\bar{X} = 35.43$	
S.D. = 4.91	
K.R. <sub>21</sub> = .58	
S.E.M. = 3.18	

Table 5

TOTAL TEST  
Score Statistics and Reliability

k = 144.5
N = 10
Range: 64.5 - 122.5 (59)
Median = 89.5
Mode = 0
$\bar{X} = 91.48$
S.D. = 17.19
K.R. <sub>21</sub> = .39
S.E.M. = 5.7

The Spearman Rank Order correlations were computed for all pairs of subtests, and between all subtests with the total test score. These correlations are presented in the matrix in Table 6, below.

Table 6  
Spearman Rho Rank Order Correlations

	PUNCT	SENT ORDER	COMP	CLOZE	TOTAL
PUNCT	1.00	---	---	---	---
SENT ORDER	.30	1.00	---	---	---
COMP	.42	.34	1.00	---	---
CLOZE	.88**	.35	.66*	1.00	---
TOTAL	.79**	.67*	.77**	.90**	1.00

\*\*  
 $p \leq .01, df = n$

\*  
 $p \leq .05, df = n$

The highest inter-test correlation was that between punctuation and cloze (.88\*\*). This suggests that these parts may be testing the same skills. It's important and interesting to note that students who demonstrated the ability to apply the syntactic function of punctuation should also perform well on a random cloze, which recent research indicates correlates significantly with structure scores. The correlation between the ability to "fill in the blanks" of a random cloze and the ability to write a short, guided composition (.66\*) was significant, although not as high as had originally been anticipated. Sentence ordering and cloze were not correlated highly enough to suggest any possible redundancy if

both were to be included in future test design.

The quite modest correlation between sentence ordering and guided composition (.34) likewise suggests that two quite different skill areas were being tested. The ability to apply discourse analysis, to recognize the function of sentence connectors and discourse markers in reconstructing a prewritten, scrambled passage was shown to be an ability somewhat apart from the more integrated skills involved in producing an original piece of writing.

Punctuation as a subskill correlated well with the total test score (.79<sup>\*\*</sup>), which suggested that this skill integrated with the other skills being tested. The remarkably high correlation between the cloze and the total test score (.90<sup>\*\*</sup>) makes a strong case for the role of the cloze in testing general writing proficiency, and certainly supports the trend to use it in diagnosing students' writing competence. A strong correlation was anticipated between composition and the total test since, in a sense, the actual writing sample provides content validity for the rest of the subtests. This expectation was satisfied with a correlation of .77<sup>\*\*</sup>. Sentence ordering did not demonstrate very high correlations with other test elements in the way punctuation and the cloze did. Nevertheless, its correlation with the total test score was quite acceptable (.67<sup>\*</sup>) - its modest correlation with punctuation and cloze, and its significant correlation with the total test was the most interesting and important result of this study.

#### RELIABILITY AND PRACTICAL APPLICATIONS

The relatively low reliability of the punctuation element was one of its least attractive features, a factor which would lead to its being discarded in a future such battery. According to the Spearman-Brown Prophecy Formula, the punctuation test would need to be lengthened to as many as 200 items in order to attain a reliability of .81. The high degree of variability in English punctuation also presented difficulties to the test designer. And finally, since so little time is generally spent in the formal study and practice of employing punctuation in a typical IEI classroom, it was felt the punctuation element may lack



content validity.

The sentence ordering element, on the other hand, demonstrated an exceptionally high degree of internal reliability. With its short length of only 27 items, the Kuder-Richardson formula 21 yielded an extremely respectable .91. By merely doubling its length, reliability could be raised to approach that level acceptable on a standardized test.

The cloze demonstrated a reasonable level of reliability, which could be raised to about .87 by increasing the number of items to 100. The comparative ease of administration and scoring of the cloze is appealing. As mentioned previously, its extremely strong correlation with the total test results is an important consideration, as well.

The reliability of the guided composition element was only .58, but this reliability could be raised to an acceptable .85 - nearly .90 by raising the number of items to 200-300. This could conceivably be done by increasing the weighting of the composition element; by eliciting more than one sample of writing per student; by increasing rater-reliability by procuring more scorers who were not familiar with the subjects of the test (as the "halo effect" was indeed discovered to be an insurmountable obstacle to objectivity).

#### CONCLUSION

A strong case has been made for the inclusion of a sentence ordering element as an objective test of writing proficiency. It is short, yet it achieves high internal consistency. It requires relatively little preparation, and provides the tester with the convenience of a discrete-point scoring system. Its relatively low correlations with other test elements, such as punctuation, composition, and cloze, indicates that it is measuring a different skill or skill set. Yet the sentence ordering element correlates highly with the total test results, and the skills it tests appear to be highly integrated with those elicited by the other elements. The content validity of the sentence ordering element is supported by current research in the field of discourse analysis, which suggests that the ability to read and write efficiently are dependent in part upon the student's application of sentence connectors and conjunctions,

his recognition of the cohesive devices which develop a logical line of thought in written form.

The high correlation between the cloze and other aspects of writing and its low correlation with sentence ordering indicate that the addition of a cloze element to an objective writing test is appropriate, and that it serves a complementary function when combined with the sentence-ordering test.

The results also indicate that the addition of a guided composition element, based upon a representative sample (2-3 paragraphs) of a student's written work, complements the objective elements, and that an optimal battery of tests needs to include an assessment of students' writing in a global, integrative manner.

## APPENDIX I

PUNCTUATION

Rewrite the following paragraph, using correct punctuation.

the getty oil field has been pumping petroleum out of southern californias kern river field since 1928 but by the mid 1960s production had declined to just 15000 barrels of oil a day it wasnt that kern river had dried up plenty of oil was left in the reservoir it was however tightly trapped in rock fissures and pores and natural pressure was no longer high enough to force it out therefore getty tried a new technique this new technique involved setting up generators in part of the field and injecting steam into the reservoir the steam heated the trapped oil making it much more fluid and forced it to flow toward production wells by last december getty had 128 steam generators in place if we had the price and the certainty of decontrol all would be go says james henry manager of reservoir engineering for arco oil and gas co in dallas texas

## APPENDIX II

SENTENCE ORDERING

Order the following sentences so that they form a logical and cohesive paragraph.

- 5 1. This stone too, and what it stands for, still exists.
- 1 2. Something remarkable happened three million years ago in Kenya.
- 7 3. There you can visit the relic, and many others like it.
- 9 4. It is a thrilling experience!
- 8 5. And when you do, you will realize that we -- modern men -- share the same genetic heritage with the hands that originally shaped this simple stone tool.
- 4 6. Thus, what was once an accident of nature became deliberate technology.
- 2 7. A human, a primitive man, picked up a water-smoothed stone.
- 6 8. It is preserved in the National Museum of Kenya in Nairobi.
- 3 9. With a few skillful strokes, he transformed it into a tool.

## APPENDIX III

CLOZE

For some reason, Richard West \_\_\_\_\_ been having great difficulty  
 (1)  
 in \_\_\_\_\_ to sleep lately. Last night, \_\_\_\_\_ example, he thought  
 (2) (3)  
 it \_\_\_\_\_ help if he went to bed earlier \_\_\_\_\_ usual, so at  
 (4) (5)  
 9:30 he \_\_\_\_\_ down, closed his eyes hopefully, \_\_\_\_\_ began  
 (6) (7)  
 counting sheep. Thinking of \_\_\_\_\_ those energetic little animals  
 (8)  
 jumping \_\_\_\_\_ fences made him feel energetic \_\_\_\_\_, so he  
 (9) (10)  
 stopped, went downstairs, \_\_\_\_\_ found the most boring book \_\_\_\_\_  
 (11) (12)  
 had. It was \_\_\_\_\_ "Home Rug Making". \_\_\_\_\_ the end of an  
 (13) (14)  
 \_\_\_\_\_ he had become quite interested \_\_\_\_\_ making rugs. He put  
 (15) (16)  
 the \_\_\_\_\_ down in desperation. Then he remembered someone \_\_\_\_\_  
 (17) (18)  
 him once that \_\_\_\_\_ you repeated the word "sleep" often \_\_\_\_\_,  
 (19) (20)  
 it would finally come. Fifteen minutes \_\_\_\_\_, the people in the  
 (21)  
 apartment above \_\_\_\_\_ tapped angrily on the floor. \_\_\_\_\_ 2  
 (22) (23)  
 o'clock he finally took \_\_\_\_\_ sleeping pill. It had \_\_\_\_\_  
 (24) (25)  
 effect. At three, he got \_\_\_\_\_ and walked around his room \_\_\_\_\_  
 (26) (27)  
 least twenty times. At four \_\_\_\_\_ did a deep-breathing exercise.  
 (28)  
 \_\_\_\_\_ five, he stared at a spot \_\_\_\_\_ the ceiling until his  
 (29) (30)  
 \_\_\_\_\_ hurt. At six his eyes \_\_\_\_\_ to feel heavy, and finally  
 (31) (32)

\_\_\_\_\_ dropped off. His alarm clock \_\_\_\_\_ at seven. This  
 (33) (34)  
 morning at \_\_\_\_\_ his boss looked at \_\_\_\_\_ gravely, shook his  
 (35) (36)  
 head, and \_\_\_\_\_ in a critical tone of \_\_\_\_\_, "See here, West!  
 (37) (38)  
 You've been \_\_\_\_\_ very tired lately. Obviously, you \_\_\_\_\_  
 (39) (40)  
 been getting \_\_\_\_\_ sleep. I really think that you \_\_\_\_\_ try  
 (41) (42)  
 going to bed earlier!"

ANSWER KEY:

- |                                     |                               |
|-------------------------------------|-------------------------------|
| (1) has                             | (22) him, his                 |
| (2) getting, going                  | (23) at                       |
| (3) for                             | (24) a                        |
| (4) could, would, might             | (25) no, little               |
| (5) than                            | (26) up                       |
| (6) lay, laid                       | (27) at                       |
| (7) and                             | (28) he                       |
| (8) all                             | (29) at                       |
| (9) over                            | (30) on                       |
| (10) too, himself, physically       | (31) head, eyes               |
| (11) and                            | (32) began, started           |
| (12) he                             | (33) he                       |
| (13) called, named, entitled, about | (34) rang                     |
| (14) at                             | (35) work                     |
| (15) hour, article                  | (36) him, Richard, West       |
| (16) in                             | (37) said                     |
| (17) book                           | (38) voice                    |
| (18) told, telling                  | (39) looking, acting, feeling |
| (19) when, if                       | (40) haven't                  |
| (20) enough                         | (41) enough                   |
| (21) later                          | (42) must, should, better     |

## APPENDIX IV

COMPOSITION ELEMENTInstructions:

Write one paragraph about your favorite month. In the paragraph, answer the following questions: --

- What is the month you like the best?
- How is the weather during this month?
- What color are the trees? What is happening to other forms of life?
- What season is your favorite month in?
- What is the general feeling of the month in your country? Here?
- What do people wear? What do children like to do during this month?
- What do you usually like to do during this month?
- Why is this your favorite month?

## REFERENCES

- Harris, David, 1968. *Testing ESL*, McGraw-Hill.
- Lado, Robert, 1961. *Language Testing*, Longman.
- Leakey, Richard, 1977. *Origins*, New York, E. P. Dutton.
- Oller, John, 1976. Language Testing Today. *English Teaching Forum* 3, pp. 22-27.
- O'Neill, R. O., 1970. *English in Situations*, London, Oxford University Press.
- Rivers, Wilga, 1963. *Teaching Foreign Language Skills*.
- Quirk, Randolph, 1972. *Grammar of Contemporary English*. Longman, p. 1055.



MODELS FOR NEW ENGLISHES

Braj B. Kachru

The use of the concept 'model' for the acquisition of English is discussed from historical, attitudinal, formal and functional perspectives, primarily with reference to the institutionalized non-native varieties of English. It is argued that the international and intranational uses of English must be studied within the sociocultural and linguistic contexts in which a variety of English is used. The understanding of such sociolinguistic profiles for each variety of English should contribute to the use of appropriate parameters for determining 'intelligibility', and for the study of the formal and functional motivations for the nativization of English. A distinction is made between acquisitional deficiency, and formal and functional difference. The former may result in 'mistakes' and the latter in 'deviations'; the deviation being an integral part of nativization. In view of the unprecedented global uses of English, a case is made for a polymodel approach for teaching English, as opposed to the monomodel approach generally recommended in literature on language pedagogy and TESL. It is claimed that a relationship must be set up between linguistic form and 'communicative units'. It is possible to set up such a relationship by establishing what J. R. Firth terms "renewal of connection" between language form and language use within the appropriate "contexts of situation" in which English is used, in native and institutionalized non-native varieties.

INTRODUCTION<sup>1</sup>

In discussing the concept "model," a distinction has to be made between the use of this term in theory construction -- for example, a model for linguistic description (see, e.g., Revzin 1966) -- and its use in pedagogical literature, where model is sometimes interrelated with method (see, e.g., Brooks 1960; Christophersen 1973; Cochran 1954; Finocchiaro 1964; Gauntlett 1957; Halliday et al., 1964; Lado 1964; and Stevick 1957). In pedagogical literature the term "model" is used in two senses: first, in the sense of acceptability, generally by the native speakers of a language; second, in the sense of fulfilling codified prerequisites according to a given "standard" or "norm" at various linguistic levels. In this sense, then, we may say that a model provides a proficiency scale. This scale may be used to ascertain if a learner has attained proficiency

according to a given norm. The term "norm" is again used in two senses: in one sense it entails prescriptivism, and in another sense, it entails conformity with the usage of the majority of native speakers, defined statistically (for a detailed discussion, see Lara 1976).

#### MOTIVATIONS FOR A MODEL

The question of a model for English has acquired immense pedagogical importance, mainly for two reasons. First, non-native varieties of English have emerged in areas such as South Asia (Kachru 1969 and later), Southeast Asia (Crewe 1977; Richards and Tay 1979), Africa (Spencer 1971a), the Philippines (Llamzon 1969), and the West Indies (Craig 1979; Haynes 1979). Second, in those areas where English is a native language, as in North America and Scotland, this question of model has often been raised with reference to bidialectism.

The identification of specific "non-standard" dialects leads to questions: Which dialect should be taught for what function? and what should be the role of bidialectism in the school system? These and related questions are being debated in educational and linguistic circles (see, e.g., Bailey 1970; Bernstein 1964; Burling 1970; Ellis 1967; Labov 1966, 1969; Riley 1978; Shuy 1971; Sledd 1969; Stewart 1970; and Wolfram 1970). Educators and linguists are also concerned about maintaining national and international intelligibility in various varieties of English (see, e.g., Christophersen 1960; Kachru 1976a; and Prator 1968).

We may discuss "model" either as a general concept, or as a language-specific concept. In language-specific terms, for example, as in the case of English, one has to discuss it in the context of sociocultural, educational, and political motivations for the spread of English. The term "spread" is used here to refer to "an increase, over time, in the proportion of a communications network that adopts a given language variety for given communicative function" (Cooper 1979:23).

The question of a "model" is then also related to the question of language spread. In the case of the spread of English, one might ask, Does English have an organized agency which undertakes the job of providing direction toward a standardized model, and toward controlling language

change-as is the case, for example, with French? Such attempts to control innovations or deviations from a "standard" in English through an Academy were not taken very seriously in Britain or in North America. The first such proposals by Jonathan Swift in Britain (around 1712) and by John Adams (in 1821: see Heath 1977) in America were not received with enthusiasm. One must then ask: In spite of the non-existence of an organized Academy, what factors have determined linguistic "etiquette" in English, and what models of acquisition have been suggested?

The documented models of English have no authority of codification from a government or a body of scholars as is the case, for example, with Spanish (see Bolinger 1975:569) or French. The sanctity of models of English stems more from social and attitudinal factors than from reasons of authority. These models, more widely violated than followed, stand more for elitism than for authority - and in that sense they have a disadvantage. The native models of English were documented partly for pragmatic and pedagogical reasons. There was a demand from the non-native learners of English for materials on learning and teaching pronunciation, for standards of usage and correctness, and for linguistic "table manners" for identifying with native speakers.

Some native speakers also wanted "authoritative" or normative codes for "proper" linguistic behavior. Of course, there have always been linguistic entrepreneurs who have catered to such demands from consumers. In 1589 Puttenham recommended that the model should be the "usual speech of the court, and that of London and the shires lying about London within 60 miles and not much above." Cooper (1687) went a step further and provided such a book for "gentlemen, ladies; merchants; tradesmen, schools and strangers," with the enticing title The English Teacher, or The Discovery of the Art of Teaching and Learning the English Tongue.

This non-authoritarian elitist prescriptivism is also found in several manuals and books on usage. A typical title, following this tradition, is The Grammarian; or The Writer and Speaker's Assistant; comprising shall and will made easy to foreigners, with instances of their misuse on the part of the natives of England. This book by J. Beattie appeared in 1838. The often quoted work on Modern English Usage by Fowler (1926) also belongs

to this tradition. (See also, e.g., Alford 1869; Baker 1770; Also relevant to this discussion are Hill 1954; Leonard 1929; Whitten 1939.)

In English when one talks of a model, the reference is usually to two well documented models, namely Received Pronunciation (RP), and General American (GA). Non-native speakers of English, even at the risk of sounding ridiculous, often aim at a close approximation of these models. The works of Daniel Jones and John S. Kenyon encouraged such attempts. What Jones's Outline of English Phonetics (1918) or English Pronouncing Dictionary (1956) did for RP, Kenyon's American Pronunciation (1924) did for GA in a restricted sense.

What type of "standard" do these pronunciation norms provide? RP as a model is about 100 years old, and it is closely associated with the English public schools. Abercrombie, in his excellent paper, considers it unique "because the public schools are themselves unique" (1951:12). Because it is acquired unconsciously, says Abercrombie, "there is no question of deliberately teaching it." The status of RP is based on social judgement, and has no official authority. The advent of broadcasting played an important role in making RP widely known. It was therefore identified with the British Broadcasting Corporation (BBC) and also termed "BBC English" (see Gimson 1970:83; Ward 1929: ch. 1 and 2). In the changed British context, Abercrombie makes three points. First, the concept of a standard pronunciation such as RP is "a bad rather than a good thing. It is an anachronism in present-day democratic society" (1951:14). Second, it provides an "accent-bar" which does not reflect the social reality of England. "The accent-bar is a little like colour-bar - to many people, on the right side of the bar, it appears eminently reasonable" (1951:15). Finally, RP does not necessarily represent "educated English," for while "those who talk RP can justly consider themselves educated, they are outnumbered these days by the undoubtedly educated people who do not talk RP" (1951:15).

The term "General American" refers to the variety of English spoken by about 90 million people in the central and western United States and in most of Canada. (See Krapp 1919; Kenyon 1924: vii, 14). In describing GA, Kenyon was not presenting a model in the same sense in which Jones had

earlier presented his. Rather, Kenyon suggests linguistic tolerance toward various American varieties of English. He is conscious of the harm done by the elitist, prescriptivist manuals for pronunciation and therefore is concerned that "we accept rules of pronunciation as authoritative without inquiry into either the validity of the rules or the fitness of their authors to promulgate them" (1924:3). The cause for such easy "judgment" or quick "advice" on matters connected with pronunciation is that people are "influenced by certain types of teaching in the schools, by the indiscriminating use of textbooks on grammar and rhetoric, by unintelligent use of the dictionary, by manuals of 'correct English,' each with its favorite (and different) shibboleth" (1924:3).

Kenyon's distaste for linguistic homogeneity is clear when he says, "Probably no intelligent person actually expects cultivated people in the South, the East, and the West to pronounce alike. Yet much criticism, or politely silent contempt, of the pronunciations of cultivated people in other localities than our own is common" (1924:5). In his view the remedy for this intolerance is the study of phonetics. A student of phonetics "soon learns not only to refrain from criticizing pronunciations that differ from his own, but to expect them and listen for them with respectful, intelligent interest."

Now, despite the arbitrariness of the above two models, one usually is asked the questions: What is a standard (or model) for English?<sup>2</sup> And, what model should be accepted? The first question is easy, and Ward (1929: 1) has given the answer in crisp words: "No one can adequately define it, because such a thing does not exist." And, in the case of English, as Strevens (1979) says, "'standard' here does not imply 'imposed,' nor yet 'of the majority.'" One interesting aspect of standard English is that in every English-using community those who habitually use only standard English are in a minority."

#### MODEL AND THE NORM

It has generally been claimed (see, e.g., Bloomfield 1933:56) that being bilingual entails having "native-like" proficiency in a language. A rigid application of this rather elusive yardstick is evident in the fast

increasing literature and growing number of texts for the teaching of English as L2. It is more evident in the structural method which followed the tenets of structural linguistics in America. Consider for example the following, which is typical of such an attitude (see Lado 1964:89):

Authentic models: Teachers can now provide authentic pronunciation models easily for their students by means of a tape recorder or a phonograph. Visitors and professional speakers can be recorded for the benefit of students, thus bringing to the class a variety of good native speakers even when the teacher does not happen to be a native speaker of the target language.

In purely pedagogical methods, with no underlying serious theoretical framework, such as the structural method developed at the Institute of Education, London,<sup>3</sup> the same ideal goal for pronunciation was propounded.

One cannot disagree that the criterion of "native-like" control is appropriate for most language-learning situations. But then, one must pause and reconsider whether such a goal for performance can be applied to the case of English in all situations. The case of English is unique because of its global spread in various linguistically and culturally pluralistic societies; its differing roles in language planning in each English-using country; and the special historical factors involved in the introduction and diffusion of English in each English-speaking country. Therefore it is rather difficult to define the concept 'norm' for various speakers of Englishes.

#### ORIGIN OF NON-NATIVE MODELS

The origin of non-native models therefore must be related to what is termed "the context of situation" - the historical context, and the educational setting. Furthermore, it should be emphasized that the question of a "model" for English did not originally arise with reference to a model for "non-native" users of English. This issue has a rather interesting history, essentially with reference to the transplanted native varieties of English. The attitude of American English users provides a fascinating and illuminating controversy on this topic, which eventually turned into a national debate (see Heath 1977; Kahane and Kahane 1977).<sup>4</sup> This national debate provides a good case study of the relationship of

political emancipation to language, and identification of language with nationalism. The controversy of the American identity of the English language has received more attention and therefore is better known, for which credit must be given to Mencken (1919). But in Britain itself there is the case of Scottish identity, and on a far-off continent, Australia, murmurs for such identity have been heard in an occasional publication.

In the case of non-native varieties, the situation is much different. There has never been a Mencken, or a Webster. The local identity for English was never related to political emancipation or national pride. On the contrary, the general idea was that, with the end of the Raj, the English language would be replaced by a native language or languages. The demand was not for an identity with English, but for abolition of English; not for nativization of English, but for its replacement. In recent years, however, the concept has been primarily discussed with reference to non-native Englishes. What do we understand by that term? The distinction between native and non-native varieties of English (Kachru 1980 and Kachru and Quirk 1980) is crucial for understanding the formal and functional characteristics of English.

In the international context, it is more realistic to consider a spectrum of Englishes which vary widely, ranging from standard native varieties to standard non-native varieties (see Kachru 1976a, 1980 and forthcoming a; and Quirk et al. 1972:13-32). The situation of English is historically and linguistically interesting and complex for several reasons. First, the number of non-native speakers of English is significant; if the current trend continues, there will soon be more non-native than native speakers of English. At present there are 266 million native speakers and 115 million non-native speakers. That is, 33.1% of English speakers are non-native users. This figure, which includes only those who are enrolled in schools, therefore does not provide the total picture. Consider the following statement of distribution:<sup>5</sup>

## Native Varieties (in millions)

British	American	Australian	Canadian	New Zealand
55	182	13	13	3

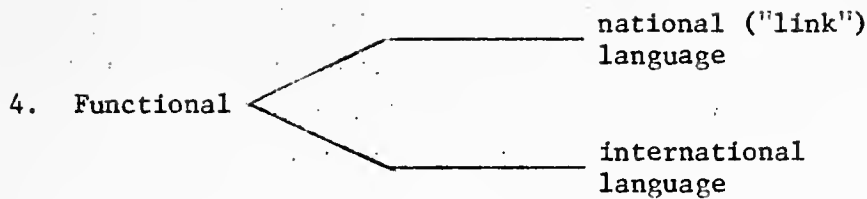
## Non-Native Varieties (in millions)

Asia (excl. USSR)	Africa	West & Cent. Europe	Soviet Union	Western Hemisphere
60	20	15	10	10

The spread of English is unique in another respect. Because the language is used in geographically, linguistically, and culturally diverse areas, its use cuts across political boundaries (Fishman et al., 1977; Smith 1980). The large range of varieties of English cannot be discussed from any one point of view. There are several, mutually non-exclusive ways to discuss their form and function. One might, for example, consider them in acquisitional terms, in sociocultural terms, in motivational terms, and in functional terms. These may further be divided as follows:

1. Acquisitional
  - first language
  - second language
  - foreign language
  
2. Sociocultural
  - transplanted
  - non-transplanted
  
3. Motivational
  - integrative
  - instrumental





A further distinction is necessary between English as a second language and English as a foreign language. (See Christophersen 1973:30-31; Quirk 1972:3-4). The second language varieties of English are essentially institutionalized varieties, as in, for example, South Asia and West Africa. The foreign language varieties are primarily performance varieties, as in Iran, Japan, etc. This distinction is also important with reference to the role and functions of English in the educational, administrative, and sociocultural context of a country in which English is used as a non-native language. The distinction between a transplanted variety (e.g., American English, Indian English) and a non-transplanted variety is important for the understanding of the acculturation and "nativization" of the transplanted varieties. (For specific case studies see, e.g., Abdulaziz 1979; Bokamba 1979; Craig 1979; Haynes 1979; Kachru 1980 and Forthcoming a; Kandiah 1979; Richards and Tay 1979; Wong 1979; and Zuengler 1979.)

In the literature, two types of motivations have been suggested for second language acquisition: integrative and instrumental. The distinction is essentially based on what function the L2 learner envisions for the acquired language. If the learner's motivation is integrative, then the desire is "to identify with the members of the other linguistic cultural group and be willing to take on very subtle aspects of their language or even their style of speech" (Prator 1968:474; his italics). On the other hand, the instrumental approach has been defined as basically "utilitarian"; a language is acquired as a linguistic tool, not as an instrument for cultural integration. Terms such as library language, auxiliary language, link language, or language for special purposes (LSP) are essentially utilitarian concepts, in which language is seen as a "restricted" code for a specific goal. In such contexts, acquiring a second culture is not the main motivation for learning the language. (See also Christophersen 1973.)

If we look at the global spectrum of English as a non-native language, we can clearly divide, as stated earlier, the non-native uses of English into two broad categories, namely, the performance varieties and the institutionalized varieties. This distinction is extremely useful and is directly related to the question of a model.

### The performance varieties

Performance varieties include essentially those varieties which are used as foreign languages. Identificational modifiers, such as Japanese English or Iranian English, are indicative of geographical or national performance characteristics. These do not indicate an institutionalized status. The performance varieties of English have a highly restricted functional range in specific contexts; for example, those of tourism, commerce, and other international transactions.

### Institutionalized varieties

It is the institutionalized varieties which have some ontological status. The main characteristics of such varieties are that (a) they have an extended range of uses in the sociolinguistic context of a nation; (b) they have an extended register and style range; (c) a process of nativization of the registers and styles has taken place, both in formal and in contextual terms; and (d) a body of nativized English literature has developed which has formal and contextual characteristics which mark it localized. On the other hand, such a body of writing is considered a part of the larger body of writing labeled English literature.

An institutionalized variety always starts as a performance variety, with various characteristics slowly giving it a different status. The main characteristics of an institutionalized variety seem to be (a) the length of time in use; (b) the extension of use;

(c) the emotional attachment of L2 users with the variety; (d) functional importance; and (e) sociolinguistic status. In the development of non-native models two processes seem to work simultaneously; the attitudinal process, and the linguistic process.

A non-native model may be treated as a competitive model for teaching English as L2 if it fulfills certain conditions. In attitudinal terms, a majority of L2 speakers should identify themselves with the modifying label which marks the non-nativeness of a model: for example, Indian English speakers, Lankan English speakers, Ghanaian English speakers. A person may be a user of Indian English in his linguistic behavior, but may not consider it the "norm" for his linguistic performance. There is thus a confusion between linguistic norm and linguistic behavior.

In linguistic terms, a viable model should describe the formal characteristics of a generally acceptable variety. If English is used in a culturally and linguistically pluralistic context, the norm for the model should cut across linguistic and cultural boundaries. It is natural that in such a variety a part of the lexicon will have been nativized in two ways. On the one hand, native items will be used in localized registers and styles to contextualize the language. On the other hand, English lexical items may have acquired extended or restricted semantic markers. This process then extends to other levels of language, as has been shown in several studies. (See, e.g., Kachru 1980, and forthcoming a.)

## DEVELOPMENT OF NON-NATIVE MODELS

The term "development" is used here not in the Darwinian sense, but in essentially a historical sense. I shall attempt to discuss it with reference to changing attitudes toward a model, in terms of a scale of acceptance. A variety may exist, but unless it is recognized and accepted as a model it does not acquire a status. A large majority of the non-native speakers of institutionalized varieties of English use a local variety of English, but when told so, they are hesitant to accept the fact.

The non-native institutionalized varieties of English seem to pass through several phases which are not mutually exclusive. At the initial stage there is a non-recognition of the local variety, and conscious identification with the native speakers. In South Asian terms, it may be called the brown sahib attitude. A "brown sahib" is more English than the Englishman; he identifies with the "white sahib" in manners, speech, and attitude, and feels that his brown or black color is a burden. At this stage an "imitation model" is elitist, powerful and perhaps politically advantageous, since it identifies a person with the rulers. This is also the stage when English is associated with the colonizer, and therefore may be a symbol of anti-nationalism.

The second stage is related to extensive diffusion of bilingualism in English, which slowly leads to the development of varieties within a variety. The tendency then is to claim that the other person is using the Indianized, Ghanaianized, or Lankanized English. The local model is still low on the attitudinal scale, though it may be widely used in various functions. South Asia provides an excellent example of this attitude. In India, for example, the norm for English was unrealistic and (worse) unavailable--the British variety. In actual performance, typical Indian English was used.

But to have one's English labeled Indian was an ego-cracking linguistic insult.

The third stage starts when the non-native variety is slowly accepted as the norm, and the division between the linguistic norm and behavior is reduced. The final stage seems to be that of recognition. This recognition may manifest itself in two ways; first in attitudinal terms, when one does not necessarily show a division between linguistic norm and linguistic behavior. This indicates linguistic realism and attitudinal identification with the variety. Only during the last twenty years or so do we find this attitude developing among the users of non-native varieties of English. Second, the teaching materials are contextualized in the native sociocultural milieu. One then begins to recognize the national uses (and importance) of English, and to consider its international uses only marginal.

The literature provides enough evidence that the institutionalized varieties of English have passed through one or more of these stages in Africa, South Asia, the West Indies, or the Philippines. I shall not elaborate on this point here.

#### FUNCTIONAL USES OF NON-NATIVE ENGLISHES

I have earlier used the term "context of situation" without explaining it in the context of the English L2 situation. There is a relationship between the context of situation, the sociolinguistic profile, and the pedagogical model. Before claiming universality for a model, one must understand that what is linguistic medicine for one geographical area may prove linguistic poison for another area.

A sociolinguistic profile should consider the type of information

suggested in Catford (1959:141-142), and in Ferguson (1966:309-315). The linguistically relevant information is as important as are the political, geographical, and economic factors. In addition, the attitudinal reactions toward an external or an internal model cannot be neglected. I shall return to that point in the two following sections.

The context of situation will then provide a cline ("a graded series") both in terms of proficiency in English, and in its functional uses. The English-using community must be seen in a new framework, in which a linguistic activity is under analysis within a specific sociocultural context. Within the framework of user and uses one has to take into consideration cline of participants, cline of roles, and cline of intelligibility.

Without the perspective of this relationship it is difficult for native speakers of English to understand the uses of non-native Englishes. This type of approach has been used and recommended in several studies. (See especially Candlin 1980; Kachru 1965, 1966, 1980 and forthcoming a and b; Richards and Tay 1979; and Richards 1979).

The institutionalized varieties of non-native English may be arranged along a lectal continuum. This continuum is not necessarily developmental but may be functional. All subvarieties within a variety (for example, basilects, mesolects and acrolects) have functional values, and may stand as clues to code diversity as well as to code development. These are however, not mutually exclusive.

Let me now briefly elaborate on the functional aspects of a cline. One can claim that, for example, in South Asia, English is used in four functions: the instrumental, the regulative, the interpersonal and the imaginative/innovative.<sup>6</sup> In each function we have a cline in performance which varies from what may be termed an "educated" or "standard" variety to a pidginized or "broken" variety. The varieties within a variety also seem to perform their functions, as they do in any native variety of English (For details see Brook 1973; Kachru 1980, especially subsection on "The Cline of Varieties"; and Quirk 1972:13-32.)

A discussion on the non-native uses of English in "un-English" contexts will entail presenting several sociolinguistic profiles relevant to a number of institutionalized varieties of English. Since in this paper I have

not set that as my goal, I will merely provide a general view of the possible functional range of non-native varieties of English.

In the case of some varieties, the English language is used in all four functions mentioned earlier. The Instrumental function is performed by English as a medium of learning at various stages in the educational system of the country. The regulative function entails use of English in those contexts in which language is used to regulate conduct; for example, the legal system and administration. The interpersonal function is performed in two senses: first, as a link language between speakers of various (often mutually unintelligible) languages and dialects in linguistically and culturally pluralistic societies; and second, by providing a code which symbolizes modernization and elitism (see Sridhar 1978.) The imaginative/innovative function refers to the use of English in various literary genres. In this function, the non-native users of English have shown great creativity in using the English language in "unEnglish" contexts. This aspect of non-native Englishes has unfortunately not attracted much attention from linguists, but has now been taken seriously by literary scholars.<sup>7</sup> (See Kachru Forthcoming b).

#### THE 'RANGE' AND 'DEPTH' OF FUNCTIONAL USES

The functional uses of the non-native varieties extend in two senses. The term "range" means the extension of English into various cultural, social, educational, and commercial contexts. The wider the range, the greater the variety of uses. By "depth" we mean the penetration of English-knowing bilingualism to various societal levels. One has to consider, for example, whether bilingualism in English is restricted to the urban upper and middle classes, or whether it has penetrated to other societal levels, too. What are the implications of these functions, and their range and depth, for a model?

The degrees of nativization of a variety of English are related to two factors: the range and depth of the functions of English in a non-native context, and the period for which the society has been exposed to bilingualism in English. The greater the number of functions and the longer the period, the more nativized is the variety. The nativization

has two manifestations, cultural and linguistic, with "cultural" here referring to the acculturation of English. The result is that, both culturally and formally, the English language comes closer to the socio-cultural context of what may be termed the adopted "context of situation." This new, changed "context of situation" contributes to the deviations from what originally might have been a linguistic "norm" or "model."

#### ATTITUDE OF NATIVE AND NON-NATIVE USERS TOWARD NON-NATIVE VARIETIES

In view of the unique developments and functions of the institutionalized non-native varieties of English, one might ask: What has been the attitude of native speakers and native users of English toward such non-native Englishes? The native speakers' attitude toward the development and the nativization of institutionalized varieties has traditionally not been one of acceptance or ontological recognition. Because of the linguistic manifestation of the nativization, these varieties have been considered deficient models of language acquisition. This attitude has not been restricted to speech performance, but extends to lexical and collocational items which are determined by the new sociocultural context in which the English language is used in Africa or Asia. It seems that the contextual dislocation (or transplantation) of English has not been recognized as a valid reason for "deviations" and innovations. Thus, the parameters for making judgments on the formal and functional uses of English continue to be culturally and linguistically ethnocentric, though the pragmatic context for such Englishes is "un-English" and "non-native" (see Kachru 1980 and Forthcoming a). Almost a decade ago, I mentioned with some elation (Kachru 1969) that with World War II a new attitude of "linguistic tolerance" had developed, which was reflected in proclamations such as "hands off pidgins" (Hall 1955), and "status for colonial Englishes." Now, almost a decade later, this statement warrants a postscript with reference to colonial Englishes. One has to qualify the earlier statement and say that this attitude was restricted to two circles. First, a body of literary scholars slowly started to recognize and accept the commonwealth literature in English written by non-native users of the language as a noteworthy lin-



guistic and literary activity. Britain was somewhat earlier in this recognition. Second, few British linguists, notably Firth (1957:97), Halliday et al. (1964), Strevens (1977:140), and Quirk (1972:26), accept the linguistic and functional distinctiveness of the institutionalized non-native varieties. It seems that even in America the linguistic fringe has been rather slow in providing such recognition and looking at these varieties in a pragmatic perspective. (For a detailed discussion, see Kachru 1976a, 1980 and Forthcoming a and b).

The non-native speakers themselves have not yet been able to accept what may be termed the "ecological validity" of their nativized or local Englishes. One would have expected such acceptance, given the acculturation and linguistic nativization of the new varieties. On the other hand, the non-native models of English (such as RP or GA) are not accepted without reservations. There is thus a case of linguistic schizophrenia, the underlying causes of which have yet to be studied. Consider, for example, the following tables. (For details see Kachru 1976a.)

TABLE 1 GRADUATE STUDENTS' ATTITUDE TOWARD VARIOUS MODELS OF ENGLISH AND RANKING OF MODELS ACCORDING TO PREFERENCE

Model	Preference		
	I	II	III
American English	5.17	13.19	21.08
British English	67.60	9.65	1.08
Indian English	22.72	17.82	10.74
I don't care		5.03	
"Good" English		1.08	

TABLE 2 FACULTY PREFERENCE FOR MODELS OF ENGLISH FOR INSTRUCTION

Model	Preference		
	I	II	III
American English	3.07	14.35	25.64
British English	66.66	13.33	1.53
Indian English	26.66	25.64	11.79
I don't know		5.12	

TABLE 3 GRADUATE STUDENTS' 'SELF LABELING' OF THE VARIETY OF THEIR ENGLISH

Identity marker	%
American English	2.58
British English	29.11
Indian English	55.64
"Mixture" of all three	2.99
I don't know	8.97
"Good" English	.27

What does such an attitude imply? In Ghana, for example, educated Ghanaian English is acceptable; but as Sey (1973:1) warns us, it does not entail competence in speaking RP since in Ghana "... the type that strives too obviously to approximate to RP is frowned upon as distasteful and pedantic". In Nigeria the situation is not different from Ghana or India (see Kachru 1976a). Bamgbose (1971:41) emphasizes that "... the aim is not to produce speakers of British Received Pronunciation (even if this were feasible!)... Many Nigerians will consider as affected or even snobbish

any Nigerians who speaks like a native speaker of English." In another English-using country, the Philippines, the model for "Standard Filipino English" is "... the type of English which educated Filipinos speak and which is acceptable in educated Filipino Circles (author's emphasis: Llamzon 1969:15). There seems to be some agreement that an external model does not suit the linguistic and sociolinguistic ecology of Africa, the Philippines, or South Asia.

#### DEVIATION, MISTAKE AND THE NORM

I have used the term "deviation" in this study, and earlier (Kachru 1965:396-398) with reference to the linguistic and contextual nativeness in the non-native varieties of English. This term needs further elucidation since it is crucial to our understanding of the question of the model. The inevitable questions concerning the linguistic and contextual deviation are: What is the distinction between a "deviation" and a "mistake"? And, how much deviation from the norm is acceptable pedagogically, linguistically and above all with reference to intelligibility.

We shall make a distinction between the terms "mistake" and "deviation" on linguistic and contextual levels. A "mistake" may be unacceptable by a native speaker since it does not belong to the linguistic "norm" of the English language; it cannot be justified with reference to the sociocultural context of a non-native variety; and it is not the result of the productive processes used in an institutionalized non-native variety of English. On the other hand, a "deviation" has the following characteristics: it is different from the norm in the sense that it is the result of the new "un-English" linguistic and cultural setting in which the English language is used; it is the result of a productive process which marks the typical variety-specific features; and it is systemic within a variety, and not idiosyncratic.

There is thus an explanation for each deviation within the context of situation. It can be shown that a large number of deviations "deviate" only with reference to an idealized norm. A number of "deviations" labeled as "mistakes" are present in native varieties of English but are not accepted when used by a non-native speaker.

In earlier studies on the non-native Englishes by educators, specialists in the teaching of English, and native speakers in general, the deviations in such varieties of English have been treated essentially as "deficiencies" in foreign language learning (e.g., Goffin 1934, Passé 1947, and Smith-Pearse 1934 for South Asian English; Hocking 1974 for African English). It seems to me that a crucial distinction is warranted between a DEFICIENT variety and a DIFFERENT variety. Deficiency refers to acquisitional and/or performance deficiency within the context in which English functions as L2. On the other hand, a different model refers to the identificational features which mark an educated variety of language distinct from another educated variety. The exponents of "difference" may be at one or more linguistic levels. The following examples from South Asian English illustrates identificational features.

### 1. Phonetics/Phonology

- (a) Series substitution involves substitution of the retroflex consonant series for the English alveolar series.
- (b) Systemic membership substitution involves the substitution of members in a system with members of another class; for example, the use of stops in place of fricative  $\theta$  and  $\delta$ , or substitution of "clear l" for "dark l."
- (c) Rhythmic interference entails the use of syllable-timed rhythm in place of the stress-timed rhythm of English (see Abercrombie 1964a:17-18; 1964b:33-34; and Kachru 1969:643).

### 2. Grammar

I shall list some characteristics discussed earlier in Kachru (1965, 1969, 1976b). A discussion on African varieties of English is available in Bokamba (1979), Bamgbose (1979), Sey (1973), and Zuengler (1979).

- (a) There is tendency to use complex sentences.
- (b) Selection restrictions are "violated" in be + ing constructions (e.g., use of hear and see in I am hearing, I am seeing).
- (c) A "deviant" pattern appears in the use of articles.
- (d) Reduplication occurs (e.g., small small things, hot hot tea).

(e) Interrogatives are formed without changing the position of subject and auxiliary items (e.g., What you would like to eat?)

### 3. Lexis

The productive processes used in lexis have been discussed, for example, in Sey (1973), Kachru (1965, 1975 and 1980) and Llamzon (1969). The term "lexis" includes here what may be termed non-native collocations (Kachru 1965:403-405). Consider, for example, turmeric ceremony, dung-wash, caste mark, police wala and lathi charge from Indian English, chewing-sponge, cover-shoulder, knocking fee, dunno drums, and bodom head from Ghanaian English.

### 4. Cohesiveness

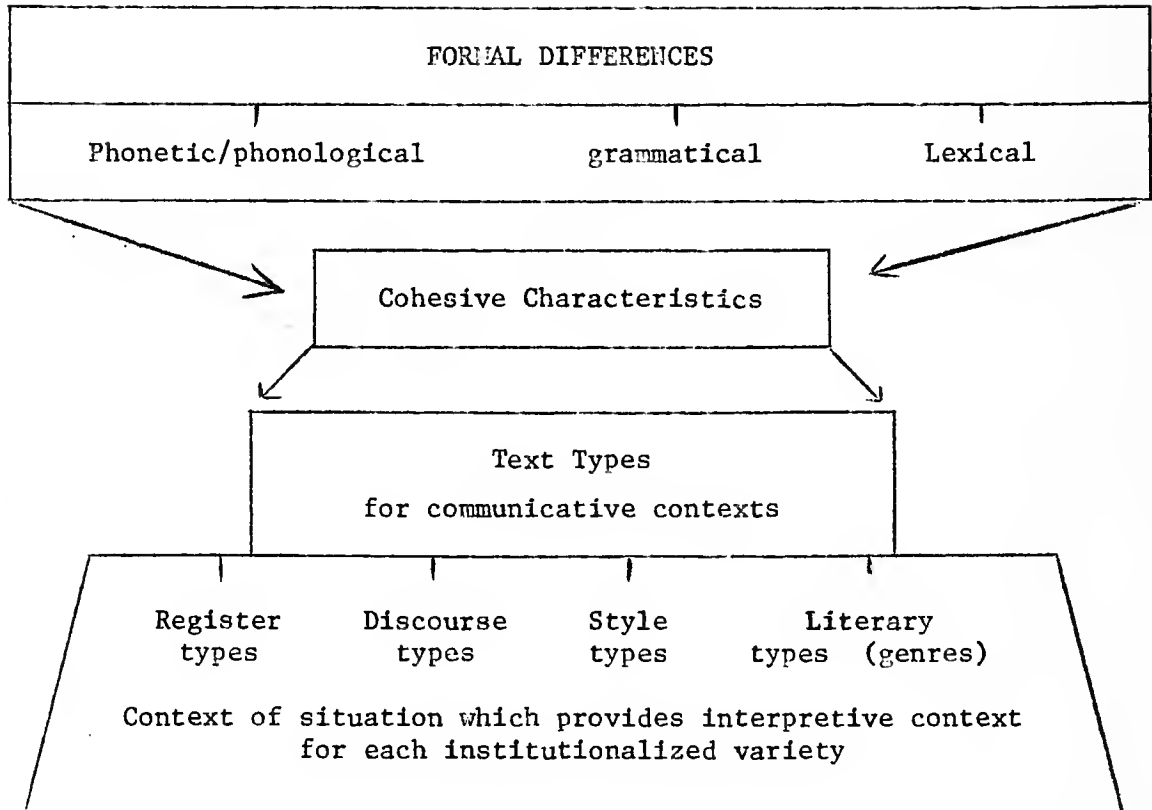
Discussion of phonology, grammar, and lexis present only one part of the total picture of the difference between "deficient" and "different" in a non-native variety. It is equally important to account for the following:

- (a) the cohesive characteristics of the text which mark it distinct, for example, in terms of its Nigerianness, Kenyanness, Indianness, or Caribbeanness;
- (b) the lexical and grammatical features which mark the register type and the style type;
- (c) the features which separate the literary genres of one non-native variety from another non-native variety.

The focus is then on setting up a relationship between the communication domains or contexts and their formal manifestations.

A non-native variety is "deviant" not only in having specific phonetic, lexical, or grammatical characteristics, but it is also "deviant" as a COMMUNICATIVE UNIT, if we compare it with other native or non-native communicative units. It is therefore necessary to establish what Firth terms a "renewal of connection" (see Firth 1956:99 and 1957:175) between the "interpretive context" ("the context of situation"), which gives the text a meaning, and its formal characteristics. The "differences" in each institutionalized non-native variety may thus be viewed in a larger context, which incorporates the "context of situation," and not purely from the

view of language deficiency. Consider the following:



If one adopts a functional view of the institutionalized varieties, it might help to abandon earlier views about two very important questions concerning intelligibility and the applicability of a monomodel approach to all the non-native varieties of English. I shall now discuss these briefly.

#### MODEL VS. INTELLIGIBILITY

In the prescriptive literature on second language acquisition, the concepts "norm" or "model" seem to play a pivotal role, primarily with regard to the non-native speaker's being "intelligible" to native speakers of English. The concept of "intelligibility" is the least researched and least understood in linguistic or pedagogical literature (see Kachru 1980; Nelson 1978). The difficulty is that intelligibility seems to have a

number of variables, and when used with reference to English it becomes more elusive. Therefore we must use the term in a specific sense. The questions one has to ask are: What is meant by intelligibility with reference to each linguistic level? Who is the judge for determining intelligibility in various varieties of English--the users of the varieties themselves, or the idealized native speakers? What parameters should be used to distinguish intelligibility between those varieties of English which are essentially regional or national (e.g., Indian English), and those varieties within a variety which have exclusively international functions? What role does a native speaker of English (and what type of native speaker) play concerning the judgment about the non-native varieties? What is the relationship between intelligibility of formal (linguistic) exponents and the contextual exponents?

'Intelligibility' has been interpreted in a rather narrow sense in earlier studies. Such studies have focused primarily on decoding a phonetic/phonological signal at the lexical level. Earlier studies, especially those of Catford (1950) and Voegelin and Harris (1951), mentioned the importance of "situation" and "effectiveness" in intelligibility. Nelson (1978) attempts to provide the parameters of intelligibility for non-native Englishes.

The intelligibility of the institutionalized non-native varieties of English forms a cline. Some speakers are more intelligible than are others, the variables being education, role, region, etc. The situation in the non-native varieties is not different from that in Britain or the U.S.A. The situation in Britain has been succinctly presented by Ward (1929:5):

It is obvious that in a country the size of the British Isles, any one speaker should be capable of understanding any other when he is talking English. At the present moment, such is not the case: a Cockney speaker would not be understood by a dialect speaker of Edinburgh or Leeds or Truro, and dialect speakers of much nearer districts than these would have difficulty in understanding each other.

In the well known cone-shaped diagram (see Ward 1929:5) Daniel Jones has graphically represented the situation: "as we near the apex, the divergencies which still exist have become so small as to be noticed only by a finely trained ear" (Ward 1929:6). Ward also rightly presents the

argument of "convenience or expediency" (1929:7) observing that "the regional dialect may suffice for those people who have no need to move from their own districts."

The case seems to be identical to that of non-native varieties of English. Intelligibility then has to be defined in regional, national, and international terms.

#### MONOMODEL VS. POLYMODEL APPROACH

In view of the special characteristics of the English speech community in various parts of the world, the pragmatic question is: Is it possible to suggest a monomodel approach, as opposed to a polymodel approach (Kachru 1977)? A monomodel approach presupposes that there is a homogeneous English L2 speech community, and that the functional roles assigned to English in each area are more or less identical. More important, it assumes that the goals for the study of English in various parts of the world are more or less similar. Such a position presupposes that the "context of situation" for the use of English in all the English-speaking areas is identical. It has already been demonstrated that such is not the case (see, e.g., Kachru 1976, 1980; Richards 1972; Strevens 1977).

The assumptions underlying a polymodel approach are diametrically opposed to the monomodel approach. A polymodel approach is based upon pragmatism and functional realism. It presupposes three types of variability in teaching English for cross-cultural communications; namely, variability related to acquisition, variability related to function, and variability related to the context of situation. We may then have to recognize a cline in terms of the formal characteristics of an L2 variety of English; functional diversity in each English speaking area; and diversity in proficiency.

The concept of "cline of bilingualism" (Kachru 1965:393-396) may, therefore, be recognized as fundamental for the discussion of a model for English. The cline applies not only to the proficiency at the phonetic/phonological levels; it must also be interpreted in a broader sense, including the overall sociolinguistic context.



## CONCLUSION

And now, in conclusion, let us face reality. The truth is that the non-native Englishes--institutionalized or non-institutionalized--are linguistic orphans in search of their parents. Several native- and non-native users of English do not understand that they are adding insult to injury by calling these varieties "bastard Englishes." The development of such varieties is not unique to English; in a lesser degree Hindi, Persian, French, and Spanish have also developed such transplanted varieties.

The problem is that even when the non-native models of English are linguistically identifiable, geographically definable, and functionally valuable, they are still not necessarily attitudinally acceptable. There is an "accent bar" which continues to segregate the non-native users. The acceptance of a model depends on its users: the users must demonstrate a solidarity, identity, and loyalty toward a language variety. In the past, the Americans demonstrated it (though not unanimously), and the result is a vigorous and dynamic American English. But then, when it comes to recognizing and accepting the varieties within American English, or accepting other non-native Englishes, Americans have shown reluctance, condescension, or indifference. The users of non-native varieties also seem to pass through linguistic schizophrenia, and cannot decide whether to accept a mythical non-native model, or to recognize the local functional model instead.

I must also mention the unique international position of English, which is perhaps unparalleled in the history of the world. For the first time a natural language has attained the status of an international (universal) language, essentially for cross-cultural communication. Whatever the reasons for the earlier spread of English, we should now consider it a positive development in the twentieth-century world context. We should realize that this new role of English puts a burden on those who use it as their first language, as well as on those who use it as their second language. This responsibility demands what may be termed "attitudinal readjustment." I have elsewhere discussed "the seven attitudinal sins" (Kachru 1976a:223-229) which the native speakers are committing in their

attitude toward the non-native varieties; a classic case is presented in Prator (1966).

The non-native users' attitudinal readjustment toward English entails the following acts, among others. First, non-native users must now dissociate English from the colonial past, and not treat it as a colonizer's linguistic tool.

Second, they must avoid regarding English as an evil influence which necessarily leads to Westernization. In South Asia and Africa the role of English in developing nationalism and mobilizing the intelligentsia at large for struggles toward freedom cannot be overemphasized. Although it is true that such use of English has resulted in a linguistic elitism, that has also been true in the past of Sanskrit and Persian, and recently of Hindi.

Third, non-native users should accept the large body of English literature written by local creative writers as part of the native literary tradition. Indian English Literature, West African literature, and Caribbean English literature not only have pan-national reading publics, but have also become part of a larger body of world writing in English. These literatures do not only interpret the national traditions and aspirations to readers across linguistically and culturally pluralistic areas. In addition, these literatures also have an international reading public. (See, e.g., for Indian English literature, Kachru 1976b:168-173, 1978a, 1978b, and Forthcoming b; Lal 1969:i-cliv; for other literatures in English see Bailey and Robinson 1973.)

Fourth, it is important to distinguish between the national and the international uses of English. It is primarily the national uses of the institutionalized varieties which contribute toward the nativization of these varieties.

Fifth, non-native users ought to develop an identity with the local model of English without feeling that it is a "deficient" model. The local (non-native) models of English are functionally as much a part of the linguistic repertoire of people as are the native (non-Western) languages. After all, in Asia or Africa it is not unusual to find that the number of users of English exceeds the number of speakers of several of what the

Indian constitution terms "scheduled languages" (or nationally recognized languages). In India, the number of English-using bilinguals is about 3% of the total population; the numbers of speakers of six scheduled languages are close to or even much less than this figure, i.e., Assamese (1.63%), Kannada (3.96%), Kashmiri (.45%), Malayalam (4%), Oriya (3.62%), and Punjabi (3%).

The international profile of the functions of English is encouraging: we may at last have a universal language as an offshoot of the colonial period. In this context, two questions may be asked: First, is there a coordinating agency which has a realistic view of the international and national functions of English? Second, do the non-native users of English feel that any significant theoretical and methodological leadership is being provided by those British or U.S. agencies which are involved in the teaching or diffusion of English? The answers to these questions, while not discussed in this paper, are closely related to our concern for studying English in the world context.

#### FOOTNOTES

<sup>1</sup>This paper is a prepublication version of an invited paper presented at the Conference on "Progress in Language Planning" organized by the Language Research Center of the William Paterson College, New Jersey, April 30 - May 1, 1979, and will appear in Progress in Language Planning edited by Juan Cobarrubias and Joshua Fishman (The Hague: Mouton, in press). I am grateful to several agencies for their support of my research on this and related topics on non-native varieties of English, specifically to the Research Board of the Graduate College and the Center for International Comparative Studies, both of the University of Illinois at Urbana-Champaign. I am also grateful to Cecil Nelson, S.N. Sridhar, and Ladislav Zgusta for their comments and suggestions on an earlier version of this paper.

<sup>2</sup>I should mention that other models, such as Scottish (English) or Australian, have been suggested in the literature. But the main viable models in the past have been RP and GA.

<sup>3</sup>The term "structural" in this method is not related to structural linguistics as understood in North America or in Britain.

- <sup>4</sup>Also see Jones (1966) for a survey of the "triumph" of English and "a history of ideas concerning the English tongue--its nature, use, and improvement--during the period 1476-1660."
- <sup>5</sup>See Gage and S. Ohannessian (1977).
- <sup>6</sup>My view of these four terms is somewhat different from that of Basil Bernstein, who originally used these terms. The functional model proposed in Halliday (1973) extends the model to nine language functions: instrumental, regulatory, interactional, personal, heuristic, imaginative, representative or informative, ludic, and ritual.
- <sup>7</sup>"This fast-growing body of writing provides impressive evidence for linguistic and contextual nativization of the English language. The result is the development of English literatures with areal modifiers, such as West African English literature, Indian English literature, Caribbean English literature, and so on. These modifiers convey not only the geographical variation, but the cultural and sociolinguistic attitudes, too. These literatures are one manifestation of the national literatures in multilingual and multicultural non-Western English-using nations. In India, for example, one can claim that there are only three languages in which pan-Indian literature is produced with an all-India reading public, English, Sanskrit, and Hindi" (Kachru 1980). For a detailed bibliography on commonwealth literature in English, specifically in Africa, India, and the West Indies, see Narasimhaiah (1976).

## REFERENCES

- Abdulaziz, M. 1976. Influence of English on Swahili: a case study of language development. Paper presented at the Conference on English in Non-Native Contexts, University of Illinois, Urbana.
- Abercrombie, D. 1951. R.P. and local accent. *The Listener*, 6 September 1951. [Reprinted in Abercrombie, D., *Studies in phonetics and linguistics*. London: Oxford University Press.]
- Alatis, J. (ed.) 1969. *Georgetown Monograph on language and linguistics*. Washington D.C.: Georgetown University Press.
- (ed.) 1976. *International dimensions of bilingual education*. Georgetown University roundtable on languages and linguistics, 1978. Washington D.C.: Georgetown University Press.
- and G. R. Tucker. (eds.) 1979. *Language in public life*. Georgetown University roundtable on languages and linguistics, 1979. Washington D.C.: Georgetown University Press.
- Alford, H. 1869. *A plea for the Queen's English*. London: Strahan.
- Avis, W. S. 1967. *A dictionary of Canadianisms on historical principles*. Toronto.
- Bailey, B. L. 1970. Some arguments against the use of dialect readers in the teaching of initial reading. *The Florida FL Reporter* spring/fall 3.47.
- Bailey, R. W. and M. Gorlach. Forthcoming. *English as a World language*.
- Baker, R. 1770. *Reflections on the English language: (1779) 2nd edition entitled 'remarks on...'*. London: bell.
- bangboşe, A. 1971. *The English language in Nigeria*. In Spencer 1971a.
- 1979 ms. *Issues in the investigation of a standard Nigerian English*.
- Bazell, C. E., et.al. (eds.), 1966. *In memory of J. P. Firth*. London Longmans.
- Beattie, J. 1838. *The grammarian; or, the writer and speaker's assistant; comprising shall and will made easy to foreigners, with instances of their misuse on the part of the natives of England*. London.
- Bernstein, B. 1964. Elaborated and restricted codes: their social origins and some consequences. *American Anthropologist* 66.55-69.
- Bloomfield, L. 1933. *Language*. New York: Holt, Rinehart and Winston.
- Bokamba, E. ms. *The Africanization of English*.
- Bolinger, D. 1975. 2nd edition. *Aspects of language*. New York: Harcourt Brace Jovanovich, Inc.
- Bright, W. (ed.) 1966. *Sociolinguistics: proceedings of the UCLA sociolinguistics conferences, 1964*. The Hague: Mouton.
- Brook, G. L. 1973. *Varieties of English*. London: Macmillan.

- Brooks, N. 1960. *Language and language learning: theory and practice.* New York: Harcourt, Brace and World.
- Burling, R. 1970. Colloquial and standard written English: some implications for teaching literacy to non-standard speakers. *The Florida FL Reporter* spring/fall 9-15,47.
- Candlin, C. 1980. Discoursal patterning and the equalising of interpretive opportunity. In Smith 1980.
- Catford, J. C. 1950. Intelligibility. *English Language Teaching* 1.7-15.
- . 1959. The teaching of English as a foreign language. In Quirk and Smith 1959.
- Christophersen, P. 1960. Toward a standard of international English. *English Language Teaching* 14.127-138.
- . 1973. *Second-language learning: myth and reality.* Baltimore: Penguin Books.
- Cochran, A. 1954. *Modern methods of teaching English as a foreign language. A guide to modern material with particular reference to the Far East.* Washington, D.C. Educational Services.
- Cooper, R. L. 1979. Language planning, language spread, and language change. In Alatis and Tucker 1979.
- Craig, D. R. 1978. Aspects of Caribbean English. Paper presented at the Conference on English in Non-Native contexts, University of Illinois, Urbana.
- Crewe, W. J. (ed.) 1977. *The English language in Singapore.* Singapore: Eastern Universities Press.
- Ellis, D. S. 1967. Speech and social status in America. *Social Forces* 45.431-437.
- Fasold, R. and R. Shuy. (eds.) 1970. *Teaching standard English in the inner city.* Washington, D.C.: Center for Applied Linguistics.
- Ferguson, C. A. 1966. National sociolinguistic profile formulas. In Bright 1966.
- Finnocchiaro, M. 1964. *English as a second language from theory to practice.* New York: Regents Publishing Co., Inc.
- Firth, J. R. 1956. Descriptive linguistics and the study of English. In Palmer 1960.
- . 1957. A synopsis of linguistic theory 1930-55. In Palmer 1968.
- Fishman, J., R. L. Cooper, A. W. Conrad. 1977. *The spread of English.* Rowley, MA: Newbury House Publishers.
- , et al. (eds.) 1968. *Language problems in developing nations.* New York: John Wiley & Sons.
- Fowler, H. W. 1926. *A dictionary of modern English usage.* London: Oxford University Press.
- Gage, W. W., and O. S. Ohannessian. 1977. ESOL enrollments throughout the world. *Linguistic Reporter* November [Reprinted in *English Teaching Forum* July 1977.]

- Gauntlett, J. O. 1957. Teaching English as a foreign language. London: Macmillan.
- Gimson, A. C. 1962. An introduction to the pronunciation of English. London: Edward Arnold.
- Goffin, R. C. 1934. Some notes on Indian English. S.P.E. Tract No. 41. Oxford.
- Hall, R. A. 1955. Hands off pidgin English! Sydney: N.S.W.
- Halliday, M. A. K. 1973. Explorations in the functions of language. London: Edward Arnold.
- , A. McIntosh, and P. Strevens. 1964. The linguistic sciences and language teaching. London: Longmans.
- Haynes, L. 1968. Caribbean English: formal and functional aspects. Paper presented at the Conference on English in Non-Native Contexts, University of Illinois, Urbana.
- Heath, S. B. 1977. A national language academy? Debate in the nation. Linguistics: an International Review 189.9-43.
- Hill, A. A. 1954. Prescriptivism and linguistics in English teaching. College English April 1954. [Reprinted in Allen (ed.), Readings in Applied English Linguistics. New York: Appleton-Century, 1958.]
- Hocking, B. D. W. 1974. All what I was taught and other mistakes. A handbook of common errors in English. Nairobi, Oxford University Press.
- Jones, D. 1916. An outline of English phonetics. [Revised edition, 1956.] Cambridge: Heffer.
- . 1956. Everyman's English pronouncing dictionary. London: Dent.
- Jones, R. F. 1965. The triumph of the English language. Stanford: Stanford University Press.
- Kachru, B. B. 1965. The Indianness in Indian English. Word 21.391-410.
- . 1966. Indian English: a study in contextualization. In Bazell 1966.
- . 1969. English in South Asia. In Sebeok 1971.
- . 1973. Toward a lexicon of Indian English. In Kachru, et al. (eds.) 1973.
- . 1975. Lexical innovations in South Asian English. International Journal of the Sociology of Language 4.55-94.
- . 1976a. Models of English for the third world: white man's linguistic burden or language pragmatics? TESOL Quarterly 10.2. 221-239.

- Kachru, B. B. 1976b. Indian English: a sociolinguistic profile of a transplanted language. In Kachru 1976c.
- (ed.) 1976c. Dimensions of bilingualism: theory and case studies. Special issue of Studies in Language Learning, Urbana: Unit for Foreign Language Study and Research, University of Illinois.
- 1977. The new Englishes and old models. English Language Forum July.
- 1978a. Toward structuring code-mixing: an Indian perspective. In Kachru and Sridhar 1978.
- 1976b. Code-mixing as a communicative strategy in India. In Alatis 1978.
- 1979. The Englishization of Hindi: language rivalry and language change. In Rauch and Carr 1979.
- 1980. The pragmatics of non-native varieties of English. In Smith 1980.
- Forthcoming a. The Indianization of English: The English Language in India. New Delhi: Oxford University Press.
- Forthcoming b. South Asian English. In Bailey and Görlach forthcoming.
- , et al. (eds.) 1973. Issues in linguistics: papers in honor of Henry and Renée Kahane. Urbana: University of Illinois Press.
- and S. N. Sridhar. (eds.) 1978. Aspects of sociolinguistics in South Asia. Special issue of International Journal of the Sociology of Language 16.
- Kahane, H. and R. Kahane. 1977. Virtues and vices in the American language: a history of attitudes. TESOL Quarterly 11.2.
- Kandiah, T. 1978. Disinherited Englishes: the case of Lankan English. Paper presented at the Conference on English in Non-Native Contexts, University of Illinois, Urbana.
- Kenyon, J. S. 1924. American pronunciation. Ann Arbor: George Wahr Publishing Company.
- and T. A. Knott. 1953. A pronouncing dictionary of American English. Springfield, Mass.: Merriam.
- Krapp, G. P. 1919. Pronunciation of standard English in America. New York: Oxford University Press.
- Lado, R. 1964. Language teaching: a scientific approach. New York: McGraw Hill.
- Labov, W. 1966. Some sources of reading problems for Negro speakers of non-standard English. NCTE Spring Institute on New Directions in Elementary English. Mimeographed 1-38.
- 1969. The logic of non-standard English. In Alatis 1969.
- Lal, P. 1969. Modern Indian poetry in English: an anthology and a credo. Calcutta, Writers Workshop.



- Lara, L. F. 1976. El concepto de norma en linguística. Mexico, D.F.: El Colegio de Mexico.
- Leonard, S. A. 1929. The doctrine of correctness in English usage, 1700-1800. University of Wisconsin Studies in Language and Literature No. 25.
- Llamzon, T. A. 1969. Standard Filipino English. Manila: Anteneo University Press.
- Mencken, H. L. 1919. The American language. New York: Alfred A. Knopf.
- Narasimhaiah, C. D. 1976. Commonwealth Literature: A handbook of select reading lists. Delhi: Oxford University Press.
- Nelson, C. L. 1978. Intelligibility in non-native varieties of English. Paper presented at the Conference on English in Non-Native Contexts, University of Illinois, Urbana.
- Palmer, F. R. (ed.) 1968. Selected papers of J. R. Firth 1952-59. London: Longmans.
- Pickering, J. 1816. A vocabulary or collection of words and phrases which have been supposed to be peculiar to the United States of America. In *The Beginnings of American English; Essays and Comments*. Chicago: University of Chicago Press, 1931.
- Platt, J. T. 1975. The Singapore English speech continuum and basilect 'singlish' as a 'creoloid'. *Anthropological linguistics* 17.7.
- 1976. The sub-varieties of Singapore English: their sociolectal and functional status. In W. Crewe 1977.
- Prator, C. H. 1968. The British heresy in TESL. In Fishman, et al. 1968.
- Puttenham, G. 1589. *Arte of English poesie*. London.
- Quirk, R., and A. H. Smith. 1959. *The teaching of English*. London: Martin Secker and Warburg. [Reprinted in *Language and Language Learning Series* 1964. London: Oxford University Press.]
- , et al. 1972. *A grammar of contemporary English*. London: Longmans.
- Rauch, I. and G. Carr (eds.). 1979. *Linguistic method: essays in honor of Herbert Penzl*. The Hague: Mouton.
- Revzin, I. I. 1966. *Models of Language*. London: Methuen. [Originally published in Russian, 1962.]
- Richards, J. and M. W. J. Tay. 1979. Norm and variability in language use. In Smith 1980.
- Riley, R. D. 1978. Should we teach urban black students standard English? *Lektos: Interdisciplinary Working Papers in Language Sciences* III.1.93-119.
- Sebeok, T. (ed.) 1971. *Current trends in linguistics*. Vol. VII. The Hague: Mouton.
- Sey, K. A. 1973. *Ghanian English: an exploratory survey*. London and Basingstoke: Macmillan.

- Shuy, R. W. 1971. Social dialects and second language learning: a case of territorial overlap. TESOL Newsletter September-December.
- Sledd, J. 1969. Re dialectalism: the linguistics of white supremacy. English Journal 58.1307-1315, 1329.
- Smith, L. (ed.) 1980. English for cross-cultural communication. London: Macmillan.
- Smith-Pearse, T. L. N. 1934. Bombay: Oxford University Press.
- Spencer, J. (ed.) 1963. Language in Africa. London: Cambridge University Press.
- . 1971a. The English language in West Africa. London: Longmans.
- . 1971b. Colonial language policies and their legacies. In T. Sebeok.
- Sridhar, K. K. 1978. English in an urban context: A South Indian case study. Paper presented at the Conference on English in Non-native Contexts, University of Illinois, Urbana.
- Stevick, E. W. 1957. Helping people learn English: a manual for teachers of English as a second language. New York: Abingden Press.
- Stewart, W. 1970. Current issues in the use of Negro dialect in the beginning reading texts. The Florida FL Reporter 3-6.
- Strevens, P. 1977. New orientations in the teaching of English. London: Oxford University Press.
- . 1979. Forms of English: an analysis of the variables. In Smith 1980.
- Voegelin, C., and Z. Harris. 1951. Determining intelligibility among dialects. Proceedings of the American Philological Society 95.3.322-329.
- Ward, I. C. 1929. The phonetics of English. Cambridge: Heffer.
- Whitten, W., and F. Whitaker. 1939. Good and bad English. London: Newnes.
- Wolfram, W. 1970. Sociolinguistic implications for educational sequencing. In R. Fasald and R. Shuy (eds.).
- Wong, I. F. H. 1979. English in Malaysia. In Smith 1980.
- Zuengler, J. E. 1979. Ms. Kenyan English.

'TRANSFER' IN 'OVERGENERALIZATION': CONTRASTIVE LINGUISTICS REVISITED

Yamuna Kachru

The issue of 'transfer' from first language in the learning of a second language has been fiercely debated in studies of second language acquisition. As the learner centered approach to language learning has gained prominence, researchers have become more and more sceptical of the role of 'transfer' or 'interference' in second language acquisition (Corder 1967). With the impact of the theory of Interlanguage (Selinker 1972), the concept of 'overgeneralization' has been invoked consistently to explain a major part of 'errors' made by a second language learner. A number of recent studies (e.g., Larsen-Freeman 1978) point out that as learning progresses, instances of 'errors' due to 'overgeneralization' increase whereas those due to 'transfer' decrease. These studies tend to discuss 'transfer' and 'overgeneralization' as discrete and mutually irreconcilable concepts. This paper examines a number of cases of so-called 'overgeneralization' to determine if in fact only overgeneralization is involved in them. It is pointed out that at least in some cases, such overgeneralization is motivated by transfer from first language. For example, the use of progressive tenses with stative verbs or present perfect with definite past time adverbs in the English of Indian speakers, no matter what their level of proficiency, is an extension of the grammar of their first language(s). The theoretical issues raised by such 'transfer in overgeneralization' are discussed with regard to their implications for second language acquisition research and contrastive linguistics.<sup>1</sup>

INTRODUCTION

Research on second language acquisition in the seventies, as compared to the sixties, underwent a fundamental change. It would not be an exaggeration to speak of the seventies as the decade of the experimental studies (e.g., the papers in Pitchie 1978, Hatch 1978 and Richards 1978). Toward the end of the sixties, certain basic changes in the thinking about second language teaching and learning had already taken place. The contrastive analysis hypothesis of the late fifties and early sixties had yielded ground to error analysis (Corder 1967 and 1971) and to interlanguage (Selinker 1972). Most experimental studies of the seventies were based upon the models of interlanguage, monitor model (Krashen 1977) and morpheme acquisition order (see Fathman

1975 and Bailey, Madden and Krashen 1974, among others).<sup>2</sup> In study after study, it was pointed out that there is a natural sequence of second language acquisition, and learners' age and first language background do not affect this natural order significantly. The basic strategy of second language learning, as of first, is overgeneralization rather than transfer (see Burt and Dulay 1974, Fathman 1975, Taylor 1975, among others). A majority of published studies thus tried to free second language acquisition from the effects of the learners' first language, and lay the ghost of first language to rest.<sup>3</sup>

Unfortunately, the ghost has refused to be laid at rest and has come to haunt second language acquisition studies again and again. Thus, for example, Hakuta and Cancino 1972 conclude:

... we conceive the order of acquisition of English grammatical morphemes as resulting from an interplay of at least two factors. One factor, consisting of variables such as frequency and salience, seems to direct the order of acquisition toward a universal order. But a second factor, transfer from the first language, modulates the order so as to produce differences between learners of different language backgrounds. (pp. 308-309)

#### THE PROBLEM

The central problem in the whole controversy regarding the role of the first language in second language acquisition seems to be the interpretation of notions such as 'transfer' and 'overgeneralization'. In studies cited earlier as well as others, researchers have assumed that transfer and overgeneralization are discrete and well-defined and it is possible to categorize errors or 'goofs' as either due to interference or overgeneralization. There seems to be very little awareness that generalization is the mechanism through which transfer of training takes place and as such, the impetus for overgeneralization may come from not only the target language patterns but also the native language experience.<sup>4</sup> I would like to argue, on the basis of data discussed in existing literature, that there is very good evidence to suggest that *a substantial part of the cases of so-called overgenerali-*

*zation are in fact cases of transfer from first language.*

#### TRANSFER OR OVERGENERALIZATION?

Dulay and Burt 1974 cites the following as reflecting typical Spanish complement structures, but argues that the structure is also typical of English:

1. I know to do all that.
2. I finish to watch TV when it's four o'clock.

The following remarks reflect the authors' bias in favor of overgeneralization:

Replacing 'know' and 'finish' with 'want', whose frequency of occurrence is undisputed, would yield a structure children produce regularly... (p. 119)

The question is not of replacing 'know' and 'finish' with 'want', the question is, do all learners of English, whether as first or second language, produce sentences such as 1 and 2? That is, is there a stage in the learning of English when 'know', 'finish' and 'want', all govern infinitival complement?<sup>5</sup> If the answer to this question is 'no', then, the explanation for 1 and 2 is obviously in 'transfer' from Spanish rather than overgeneralization of the English pattern of 'want'. An examination of the existing literature on first language acquisition fails to reveal any overgeneralization of the infinitival complement structure in English. Even three year old English speaking children are reported as producing sentences such as the following (Clark and Clark 1977):

3. When I've cut something, I know how to flatten it again.  
(Table 9-7, p. 359)
4. Ask me if I not made mistake.  
(Table 9-4, p. 349)
5. I don't want you read that book.
6. I guess she is sick.  
(pp. 360-361)
7. I promised you you could do a somersault.  
(p. 362)

Note that in 5, 6 and 7, the complementizers to and that have been left

out, but in 3 and 4, the child seems to select a proper question word complement.

Similarly, Jain 1974 quotes the following from Indian English as instances of overgeneralization:

8. I know swimming.
9. God will ask him that why he did not make any use of his talent.
10. He knows well that what a blind man can do.
11. His behavior indicates that whom he likes more.

It can equally be argued that the above are due to transfer from the first language if the following facts are taken into account. In South Asian languages such as Hindi and Urdu, there are only two devices to mark the complement structure. One is the complementizer ki which is equivalent to the English that, the other is kā - nā 'poss - to' which is equivalent to both 'for - to' and 'poss - ing' of English. In Hindi, the infinitival complement (i.e., kā - nā) functions as a noun phrase, similar to the poss - ing of English. Verbs such as learn, know, want in Hindi govern this complement. As such, 8 is a direct translation equivalent of the Hindi sentence given below:

12. maĩ tairnā jāntā hū̃.

I to swim know

I know how to swim.

Also, the complementizer ki 'that' must occur to mark the finite complement clause as subordinate, regardless of whether it is interrogative or not. As such, 9 - 11 above are direct translation equivalents of the parallel Hindi sentences. For instance, the Hindi sentence that is equivalent to 10 is as follows:

13. vah acchī tarah jāntā hai ki ek andhā ādmī kyā kar  
he well knows that a blind man what do

saktā hai.

can

He knows well what a blind man can do.

## EVIDENCE FOR TRANSFER

Consider further the evidence for claiming that overgeneralization is inspired by the native language patterns. Schachter 1974 discusses the difficulty some language speakers have with English relative clauses. They produce relative clauses such as the following:

14. The problem that a tourist guide must solve them are numerous. Note that sentences such as 14 are common in the speech of Arabic, Chinese, Hebrew and Persian learners of English. In contrast, Indian speakers of English do not produce such sentences. The explanation for this phenomenon may be as follows. In Indian languages such as Hindi, Kannada, and Urdu, the strategies for relative clause formation do not utilize the device of leaving a pronominal copy of the relativized noun behind. In Arabic, Chinese, etc., this device is exploited to a great extent. Typologically speaking, the pronominal copy strategy of relative clause formation is more transparent in that it preserves the information about the function of the relativized noun. Compared to the relativization strategy of English, the pronominal copy strategy is thus less-marked.<sup>7</sup> No wonder this strategy is overgeneralized by the speakers of Arabic, Chinese, Hebrew, Persian and other such languages.

Next, consider the case of the use of English tense forms as in 15:

15. I have written a letter yesterday.

As the analyses of English present perfect in McCawley 1971 and Comrie 1976 show, it is more complex than comparable tense forms in Spanish, French, Hindi and other languages. Part of the complexity results from the fact that in English, the definite time temporal adverbs such as yesterday refer to the reference point rather than to the event, hence the ungrammaticality of 15 and grammatical well-formedness of 16:

16. I have written to him recently.

In Hindi, as shown in Hackman 1976, the definite time temporal adverb refers to the event, hence the translation equivalent of 15 is well-formed:

17. maĩ ne use kal ciṭṭhĩ likhĩ hai.

I ag. him yesterday letter written have

\*I have written to him yesterday.

No wonder a Hindi speaker fails to realize there is anything wrong with his English sentences such as 15.

Similarly, the grammatical features of stative vs. non-stative are a matter of grammatical construction rather than of verbal lexicon in most Indian languages. As such, it is natural for an Indian speaker of English to treat see, hear, desire, etc. as active and use them in the progressive. Consider the following from Jain 1974.

18. I was seeing all this happen in front of my house.

19. I am having a very heavy work-load this semester.

20. All these days, I have been desiring to meet you.

In most Indian languages, there are dative-subject constructions for the stative, and nominative-subject constructions for the non-stative, e.g.:

21. mujhko sab kuch dikhāĩ de rahā hai.

me to everything visible give ing is

I see everything.

22. maĩ sab kuch dekh rahā hũ.

I everything see ing am

I am looking at everything.

Notice that both the dative and nominative constructions above are in the progressive. Again, it is hard for an Indian speaker of English to realize intuitively that there is anything wrong with sentences such as 18-20.

#### CONCLUSION

Facts such as the above are hard to explain away on the basis of overgeneralization of target language patterns. In fact, so far, we have worked with simplistic notions of both overgeneralization and transfer. Just as it is too simplistic to assume that every native language feature is automatically transferred to a second language learning situation, it is too simplistic to assume that once we have



gone through the process of learning of one language, the resultant conceptual framework is simply kept rigidly separated when faced with the task of learning subsequent languages. It is more reasonable to assume that the hypotheses that worked once would be pulled out again and either suitably modified or totally rejected only on the basis of subsequent experience.

Notice that unless *overgeneralization on the basis of transfer from first language* is recognized as a valid strategy, it is difficult to account for the nativization processes that have been discussed in characterizing 'new' Englishes (Bokamba [forthcoming], B. Kachru 1965, 1980 and [in press]). The studies on international, institutionalized varieties of English show clearly that in many cases, overgeneralization on the basis of transfer from native language(s) is not due to ignorance of rule restrictions, on the contrary, it is motivated by very real linguistic needs and attitudes of the community that uses these varieties.

It seems to me that a fruitful area of research would be as follows. In order to explain the systematicity as well as variability of the interlanguage of the learner, researchers could look at the rapidly growing body of literature on linguistic universals and typological universals. It would be interesting to find out to what extent interlanguages conform to our theoretical notions regarding the nature of human languages. Also, it would be interesting to find out to what extent interlanguages exhibit typological differentiation depending upon the native language background. The few preliminary investigations that we have reports of clearly show that the typological characteristics of first languages affect the learning of second language in interesting and predictable ways (e.g., Schachter 1974, Eckman 1977).

A great many insights have been gained in the areas of putative linguistic universals as well as typological universals by both implicit and explicit contrastive analyses of natural languages (Ross 1967, Perlmutter 1971, Keenan and Comrie 1977, among others). To quote Ferguson 1968:

Contrastive analysis is basic to all linguistics since only by this approach can a general theory of language ('language universals') be constructed and only with at least implicit contrastive analysis can a particular language be fully characterized... (p. 101)

It would be a pity to throw out contrastive analysis without seriously considering its potential value in accounting for second language acquisition.

\* \* \* \* \*

#### FOOTNOTES

- <sup>1</sup>An earlier version of this paper, under the same title, was presented at the TESOL Convention, 1980, on March 6. I am grateful to J Ronayne Cowan and S. N. Sridhar for their valuable comments.
- <sup>2</sup>Such studies were based upon the 'morpheme acquisition order' discussed in Brown 1973 which is a study of first language acquisition.
- <sup>3</sup>See, however, Rosansky 1976 for a critique of such experimental studies. Some of the studies cited here (e.g., Burt and Dulay 1974) use the Bilingual Syntax Measure to collect data. Oller 1976 notes that the reliability coefficients for the various parts of the BSM fall well below the levels which are normally accepted as indicating that the test is a useful measuring instrument. Porter 1977 also points out the problems of using BSM for data collection (see also the rejoinder in Krashen 1978). Note that Taylor 1975 is cautious in stating that transfer is much more pronounced at the elementary stages of language learning. There are, however, a large number of unexplained 'errors' in the study and no discussion is provided for the variation in data. In short, the studies mentioned above and others similar in nature lead to serious doubts as to whether errors could unambiguously be characterized as due to overgeneralization.
- <sup>4</sup>An additional problem is that of procedural confusion in isolating two overlapping second language acquisition strategies: Overgeneralization and ignorance of rule restrictions. As any case of the latter can be called an instance of overgeneralization, it is not clear how one can be sure whether some phenomenon is due to overgeneralization or ignorance of rule restrictions.

- <sup>5</sup>The notion 'govern' explains facts such as the following: Certain categories of lexical items (such as verbs, adjectives, prepositions) allow certain grammatical processes whereas others block them. For example, verbs such as read, hit, believe allow passivization in English whereas verbs such as have, cost, resemble block this process. In English, verbs such as say govern full clausal complements with that, verbs such as enjoy govern only poss-ing complements, and verbs such as want govern only the infinitival complement. There are, of course, verbs that govern more than one type of complement structure.
- <sup>6</sup>The information and sentences quoted are from Clark and Clark 1977 which summarizes many first language acquisition studies. A review of primary sources failed to reveal any data or observation that would even partially answer the question about first language learners' overgeneralizing the infinitival complement in English.
- <sup>7</sup>Schachter 1974 is basically concerned with the controversy regarding 'CA as Predictive' vs. 'CA as Explanatory' hypothesis, or the strong vs. weak versions of CA. The basic argument of the study is that the data supports the predictive CA hypothesis in that due to complex differences between the English vs. Chinese-Japanese relative clause formation, the speakers of the latter languages produce less relative clauses in English as compared to the speakers of Arabic and Persian. Thus, avoidance can be said to be a subtle form of interference. In so far as the errors of the Arabic and Persian speakers are concerned, the notion of 'markedness' provides an explanation as suggested in Eckman 1977.

The term 'markedness' has been used in several senses in linguistic literature. In one sense, the more complex a certain phenomenon is, the more marked it is said to be. For instance, plural forms of nouns in English are more marked as compared to their singular counterparts, since the plural forms are more complex, i.e., they require either suffixation or internal modification as in boys, men, etc. In another sense, the less transparent a structure is, the more marked it is, e.g., (i) is more marked as compared to (ii) or (iii):

- (i) I saw the boy mowing the lawn.
- (ii) I saw the boy who was mowing the lawn.
- (iii) I saw the boy while he was mowing the lawn.

In (i), the relationship of mowing the lawn to the rest of the clause is not clear, in (ii) and (iii), the relationship is transparent. In this sense of markedness, the NP-complement structure of English is more marked as compared to languages such as Persian and Hindi.

- <sup>8</sup>Explicit contrastive analyses, of course, could be made more insightful by incorporating semantic and pragmatic factors as suggested in Kachru 1975 and 1976, and typological markedness as suggested in Eckman 1977.

## REFERENCES

- Bailey, Nathalie, Carolyn Madden and Stephen D. Krashen. 1974. Is there a natural sequence in adult second language learning? *LL* 24.2. 235-243.
- Bokamba, Eyamba. (forthcoming). The Africanization of English.
- Brown, Roger. 1973. *A first language: The early stages*. Harvard University Press.
- Clark, Herbert H. and Eve V. Clark. 1977. *Psychology and language: An introduction to psycholinguistics*. Harcourt.
- Comrie, B. 1976. *Aspect*. Cambridge University Press.
- Corder, Stephen P. 1967. The significance of learners' errors. *IRAL* V.4. Reprinted in Richards 1974. 19-27.
- Corder, Stephen P. 1971. Idiosyncratic dialects and error analysis. *IRAL* IX.2. Reprinted in Richards 1974. 158-171.
- Dulay, Heidi C. and Marina K. Burt. 1974. You can't learn without goofing. Richards 1974. 95-123.
- Eckman, Fred R. 1977. Markedness and the contrastive analysis hypothesis. *LL* 27.2. 315-330.
- Fathman, Ann K. 1975. Language background, age and the order of acquisition of English structures. *On TESOL '75*. 33-43.
- Ferguson, Charles A. 1968. Contrastive Analysis and language development. James E. Alatis (ed.) *Contrastive Linguistics and its pedagogical implications*. Monograph series on languages and linguistics, 21. Georgetown University Press, 101-112.
- Hackman, Geoffrey J. 1976. *An integrated analysis of Hindi tense and aspect system*. Ph.D. dissertation, University of Illinois (unpublished).
- Hakuta, K. and H. Cancino. 1977. Trends in second language acquisition. *Harvard Educational Review* 47. 294-316.
- Hatch, Evelyn M. (ed.) 1978. *Second language acquisition: A book of Readings*. Newbury.
- Jain, M. P. 1974. Error analysis: Source, cause and significance. Richards 1974. 189-215.
- Kachru, Braj B. 1965. The Indianness in Indian English. *Word* 21. 391-410.
- Kachru, Braj B. 1980. The pragmatics of non-native varieties of English. Larry Smith (ed.) *English for cross-cultural communication*. London, Macmillan.
- Kachru, Braj B. (in press a). South Asian English. R. W. Bailey and M. Gorlach (eds.) *English as a world language*.

- Kachru, Braj B. (in press b). *The Indianization of English: The English language in India*. New Delhi, Oxford University Press.
- Kachru, Yamuna. 1975. The semantics and syntax of the causative construction in English and South Asian languages: Implications for teaching English as a second language. ERIC document ED119 528.
- Kachru, Yamuna. 1976. Defining 'equivalence' in contrastive analysis: causative constructions in English and Hindi-Urdu. *CIEFL Bulletin* 12. 1-13.
- Keenan, E. and B. Comrie. 1977. Noun phrase accessibility and universal grammar. *LI* 8.1. 63-100.
- Krashen, Stephen D. 1977. The Monitor model for adult second language performance. M. Burt, H. Dulay and M. Finnochiaro (eds.). *Viewpoints on English as a second language*. New York, Regents. 152-161.
- Krashen, Stephen D. 1978. Is the "Natural Order" an artifact of the Bilingual Syntax Measure? *LL* 28.1. 187-191.
- Larsen-Freeman, Diane E. 1978. Evidence of the need for a second language acquisition index of development. Ritchie 1978. 127-136.
- McCawley, James. 1971. Tense and time reference in English. C. J. Fillmore and D. T. Langendoen (eds.). 1971. *Studies in linguistic semantics*. Holt. 97-114.
- Oller, John W., Jr. 1976. The measurement of bilingualism. Review essay. *MLJ* LX.7. 399-400.
- Perlmutter, David M. 1971. *Deep and surface structure constraints in syntax*. Holt.
- Porter, R. 1977. A cross-sectional study of morpheme acquisition in first language learners. *LL* 27. 47-52.
- Richards, Jack C. (ed.). 1974. *Error Analysis: Perspectives on second language acquisition*. Longman.
- Richards, Jack C. (ed.). 1978. *Understanding second and foreign language learning: Issues and approaches*. Newbury.
- Rosansky, Ellen J. 1976. Methods and morphemes in second language acquisition research. *LL* 26.2. 409-426.
- Ross, John R. 1967. *Constraints on variables in syntax*. Ph.D. dissertation, M.I.T. (unpublished).
- Ritchie, William C. (ed.). 1978. *Second language acquisition research*. Academic Press.

- Schachter, Jacqueline. 1974. An error in error analysis. *LL* 24.2. 205-214.
- Selinker, Larry J. 1972. Interlanguage. *IPAL* X.3. Reprinted in Richards 1974. 31-54.
- Taylor, Barry P. 1975. The use of overgeneralization and transfer learning strategies by elementary and intermediate university students learning ESL. *On TESOL '75*, 55-69.

PHONOLOGICAL AND MORPHOLOGICAL CONDITIONING OF {Z}  
IN THE SPEECH OF JAPANESE ESL STUDENTS

Susan Osuch-Hatzivramidis and Lonna J. Dickerson

This study examines the use of the noun plural morpheme {Z} by three Japanese speakers of English. The performance data is considered from two different points of view: (1) the long/short approach which requires the learner to attach a plural ending, either long (requiring an extra syllable) or short (not requiring an extra syllable), and (2) the traditional three-way approach which requires phonetically accurate pronunciation for the three endings /əz/, /s/, and /z/.

The results indicate a systematic patterning of the data for both approaches. For the two-way, long/short distinction, each subject has less difficulty with the long ending than with the short ending; for the three-way distinction, each subject shows a similar sensitivity to various environmental constraints, such as [voice], [sonorant], [coronal], which control phonetic accuracy in the production of the three noun plural endings. In addition, certain pedagogical implications derived from this study are discussed.

INTRODUCTION

For ESL teachers, one area of concern is the teaching and learning of the noun plural morpheme {Z}. As they observe their students struggling for mastery of this morpheme, teachers are often perplexed by questions such as: how do my students go about learning our plural system; that is, what process or processes are involved? How can I explain what appears to be a hodgepodge of random mistakes? How can I measure progress with some degree of accuracy? How can I structure my teaching of plurals for maximum learning efficiency? In response to questions such as these, this paper examines the performance of three Japanese speakers as they use the regular noun plural morpheme {Z}. The general aim of the study is to discover the systematic influences underlying their performance, thus providing insights which can be of practical value to the ESL teacher.

As shown in Figure 1 below, there are two ways to categorize the regular allomorphs of the noun plural ending. One way is to distinguish between the long ending (requiring an additional syllable for pluralization)

and the short ending (requiring no extra syllable for pluralization). For this approach, the exact phonetic rendering of the ending is unimportant. That is, if the long ending of *pages*, for example, is pronounced as either [əz] or [əs], it is counted as correct, or if the short ending of *pies*, for example, is pronounced as either [s] or [z], it also is judged correct. A second way to categorize these same endings is to make a three-way distinction: /s/ is added to nouns ending in voiceless nonsibilant phones (e.g., *tops*); /z/ is added to nouns ending in voiced nonsibilant phones (e.g., *rugs*); and /əz/ is added to nouns ending in sibilant phones (e.g., *churches*). For this approach, a more precise pronunciation is required. For example, the ending on *boys* is correct only when pronounced as [z]; it is incorrect when pronounced as [s].

Noun Plurals	Two-Way Distinction	Three-Way Distinction
<i>tops, hats, cakes</i>	SHORT	/s/
<i>rugs, boys, beds</i>		/z/
<i>churches, pages, dresses</i>	LONG	/əz/

Fig. 1. Two Ways to Categorize Regular Noun Plural Endings

## THE STUDY

### Purpose

The specific purpose of this study was to address the following issues:

1. Which of the English plural endings, long or short, is most difficult for the Japanese speakers studied? Do all the subjects share the same areas of difficulty? How are the subjects alike? How are they different?
2. How does phonological environment influence the type of plural ending used? That is, how does the specific final sound in a word affect the accuracy of the plural form added to that word?



## Subjects

The subjects in this study were three Japanese speakers, two women and a man between the ages of twenty-five and thirty-five. All three subjects placed at an intermediate level of English instruction based on the *University of Illinois English Placement Test* and all three belonged to the same class in the Intensive English Institute at the University of Illinois at Urbana-Champaign. Thus, the selection of subjects was controlled mainly for language background and proficiency level.<sup>1</sup>

## Background

A brief look at some of the past studies in English plural formation provides an important context for this study. Heidi Dulay and Marina Burt (1974) studied how 115 Spanish and Chinese speakers used eleven English morphemes, including the long and short plural forms. Using group scoring methods, they found that their subjects had more difficulty with the long plural than with the short plural. Similar results have been obtained by Larsen-Freeman (1975) for speakers of Arabic, Japanese, Persian, and Spanish. Studies of the ordering of plural endings among native English-speaking children, such as the study by Natalacio and Natalacio (1971), have also found that the short plural forms are produced more easily than the long plural form. It has thus been suggested that both native and nonnative English speakers find the short plural endings easier to produce.

Even though the studies cited above are in agreement, the need for more research is apparent. Most of the previous findings are based on the collection of small amounts of data (few test items) from large groups of subjects. Such data collection techniques are apt to result in two potential problems: (1) Data pools comprised of so few items do not always provide a representative sample of the many environments in which long and short plural morphemes occur.<sup>2</sup> (2) The inclusion of so few items encourages an analytical procedure which can be misleading. That is, past researchers, operating on the assumption that learners from the same language background and with the same general level of proficiency are alike in performance, have tended to study an entire group of subjects as a single unit. Because the individual language learner is not studied, individual variation is often obscured and, consequently, systematicity

is difficult, if not impossible, to observe except at the most general level. To correct for these problems, the present study focuses on the analysis of a large amount of data carefully controlled to include an equal number of the many environments in which the endings occur, and it also examines the performance of the language learner as an individual before pooling his performance data with the group.

Still other reasons for doing further research on noun plurals concern the scope of previous studies. By simply comparing "long" to "short" production, most of these studies have paid little or no attention to the more precise pronunciation of the three plural endings. Furthermore, by ignoring the effect on the plural endings created by different phonetic characteristics of the stem-final segment, it has been impossible to uncover many of the intricacies of the system. For this reason, this present study deals with the specific phonetic rendering of the three forms, /əz/, /s/, and /z/, as well as the more general production of long versus short endings, and presents a detailed analysis of conditioning effects of the stem-final environment.

#### Test Instrument and Analysis Procedure

The test instrument consisted of 185 pictures of objects which were illustrated in both the singular and the plural. For each stem-final consonant sound (e.g., the /t/ of *hat*, *cat*), a minimum of five pictures was used.<sup>3</sup> Fifteen items ended in a vowel sound (e.g., the vowel /ɔy/ of *boys*, the vowel /ay/ of *pies*).

The test was given orally in an informal setting; all responses were tape recorded. In each session, the same examiner conducted the test with one subject at a time. A series of randomly-ordered pictures was presented, each with a single object depicted. The examiner modelled the singular form of the object orally. Then, pictures of the same items were shown where two or more of the objects were present. The examiner asked, "What are these?", and the subject responded with the plural form of the item. The examiner never cued the subject with a plural form. If a subject forgot the name of an object, the examiner cued him or her by giving the name of the item in the singular.

For the data analysis, all responses were transcribed phonetically

from the tapes. Performance for each specific environment (e.g., endings after /m/, endings after /z/) was analyzed separately; then similar environments (e.g., endings after nasals, endings after vowels) were grouped and analyzed. The performance of each subject was first analyzed individually; then the data for the three subjects was grouped.

## RESULTS AND DISCUSSION

The results of this study can best be discussed in relation to the questions posed at the outset. Question One asked:

Which of the plural endings, long or short, is most difficult for the Japanese speakers studied? Do all the subjects have the same areas of difficulty? How are the subjects alike? How are they different?

Figure 2 below displays the areas and order of difficulty for Subject A. This summary chart indicates not only Subject A's success in forming long and short plurals for a given category of sounds (rows 1-6), but it also indicates the relative difficulty of plural formation for each sound category, as indicated by the percentages listed in the right-hand column.<sup>4</sup> For example, Subject A had no difficulty forming noun plurals after nasal and liquid sounds (row 1, 100%), only a little difficulty after sibilants and vowels (row 2, 96.5%; row 3, 95.5%), moderate difficulty after labiodental fricatives and voiced plosives (row 4, 81.8%; row 5, 77.3%), and much difficulty after voiceless plosives (row 6, 53.6%). When comparing success in producing long versus short plurals, long plurals are easier than short plurals with one exception. The short plural after nasals and liquids was easier (100% correct production) than the long plural after sibilants (96.5% correct production).

	Correct Plural Ending	Environments	Example	$\frac{\# \text{Correct}}{\# \text{Possible}}$	% Correct
1.	SHORT	NASALS /m,n,ŋ/ LIQUIDS /r,l/	<i>phones</i>	42/42	100.0%
2.	LONG	SIBILANTS /s,z,š,ž,č,ǰ/	<i>dresses</i>	55/57	96.5%
3.	SHORT	VOWELS	<i>boys</i>	21/22	95.5%
4.	SHORT	LABIODENTAL FRICATIVES /f,v/	<i>gloves</i>	9/11	81.8%
5.	SHORT	VOICED PLOSIVES /b,d,g/	<i>beds</i>	17/22	77.3%
6.	SHORT	VOICELESS PLOSIVES /p,t,k/	<i>cats</i>	15/28	53.6%

Fig. 2. Subject A. Production Classified by Type of Final Stem Sound of Singular Noun

The performance of Subjects B and C was strikingly similar to that of Subject A. For example, all three subjects used the correct plural ending 100% of the time for nouns ending in nasals or liquids; nouns ending in sibilants or vowels came next where the correct plural form occurred more than 90% of the time. Although exact percentages for each category in Figure 2 varied from subject to subject, the relative ranking of the categories was very similar.<sup>5</sup>

In answer to Question One posed earlier, Figure 3 summarizes the findings for all three subjects. It shows that not only was the long plural ending easier than the short plural ending for the group as a whole (long = 95%, short = 87%), but the pattern was identical for each of the three subjects when considered individually. While individual variation did exist, that variation was minor and it did not obscure the clear pattern.

SUBJECT	LONG		SHORT	
	# Correct	% Correct	# Correct	% Correct
A	(55/57)	97%	(104/121)	84%
B	(53/59)	90%	(107/123)	87%
C	(59/60)	93%	(112/126)	89%
Group Means		95%		87%

Fig. 3. Percentages of Correct Production for Long and Short Plural Forms

The most obvious explanation for the Japanese subjects' greater success with long plurals is language transfer. Since the Japanese language has basically a CV syllable structure, it is reasonable to assume that Japanese learners of English would have less difficulty producing the long endings, which are similar to the familiar Japanese CV sequences, than the short endings, most of which form consonant clusters.

The relative ease with which the Japanese subjects produced long plurals contrasts sharply with the relative difficulty subjects in other studies had with the same ending. As noted earlier, other researchers have observed that nonative English speakers found the short plural form to be less difficult than the long plural form. The results of this present study, then, call into serious question the claim that the short plural is universally easier than the long plural. The contradictory findings may be explained by some of the methodological problems discussed above under 'Background,' or there may be a genuine difference in subjects from different language backgrounds. Needed is further research using different types of test data, more subjects, subjects from different language backgrounds, and subjects at different levels of proficiency.

Thus far our concern has been whether or not an appropriate noun plural ending--long or short--appeared in the appropriate place. This evaluation of long and short endings was done without regard to the more exact phonetic rendering of the endings as entailed in a three-way dis-

inction. Although some patterns of ease and difficulty were apparent, the following more detailed analysis will show that the long/short approach to plural endings cannot provide a complete picture of the systematic nature of the data. In fact, the long/short approach used alone obscures many important regularities which give us a great deal of precise information about processes controlling Japanese speakers' success with the {Z} morpheme.

Let us now turn to the second question posed at the outset of this paper:

How does phonological environment influence the type of plural ending used? That is, how does the specific final sound in a word affect the accuracy of the plural form added to that word?

To answer this question we need to reexamine the data, this time evaluating the phonetic accuracy of the noun plural endings. This means that to be counted as correct, a required /s/ ending must be pronounced [s], a required /z/ ending must be pronounced [z], and a required /əz/ ending must be pronounced [əz].

For each of the three subjects, Figure 4 gives the percentages of correct production for the three endings. When we compare the percentages in Columns II and III with those in Column I, we see that the /əz/ ending is easier for all three subjects than the /s/ or /z/ endings. The status of /əz/ in this analysis parallels the status of /əz/ in the two-way analysis; all three subjects had more target production for long than for short endings. To summarize, then, whether we evaluate the subjects' performance according to the two-way long/short distinction (without regard to the exact phonetic rendering of the endings) or according to the three-way distinction (which required phonetic accuracy), the long (or /əz/) ending is easier for all three subjects.

	I	II	III
SUBJECT	Sibilant + /əz/	C <sub>vl</sub> + /s/	Vl + /z/ C <sub>vd</sub> + /z/
A ———	68%	59%	34%
B ·····	77%	69%	66%
C - - - -	70%	32%	13%

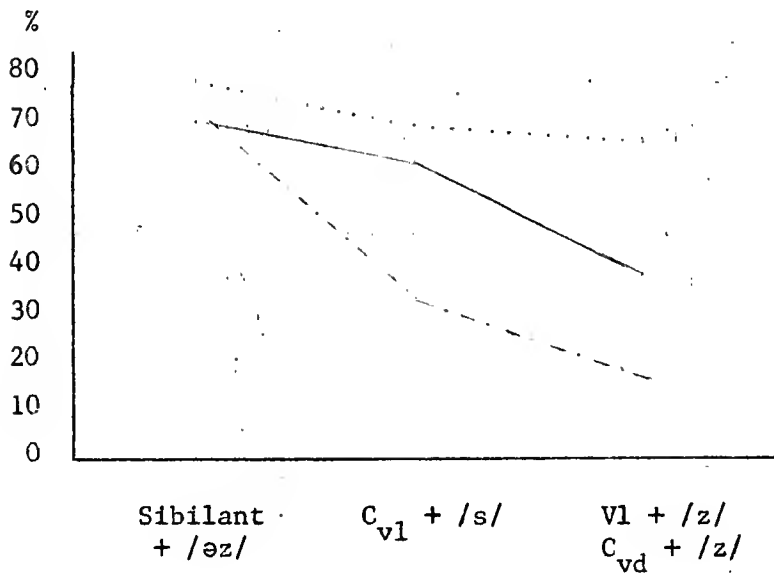


Fig. 4. Percentages and Line Graph of Correct Production for /əz/, /s/, and /z/

Observe that the /s/ ending in Column II is easier for all subjects to articulate correctly than the /z/ ending in Column III. In fact, when we compare pairs of voiceless and voiced final-stem consonants (e.g., /p/ vs. /b/, /f/ vs. /v/), a striking pattern emerges. For every pair, every subject produced an equal or greater percentage of correct {Z} after the voiceless segment when compared with the percentage of correct production after its voiced counterpart. This patterning can be explained by language transfer. Since Japanese does not have final voiced consonants,

final voicing in English is difficult, particularly when combined with the simultaneous difficulty of producing a consonant cluster.

In addition to the patterning along voicing lines, other regularities can be observed when the data is grouped according to other environmental constraints. For each of the three subjects, Figure 5 gives the percentage of correct production of /s/ and /z/ endings, excluding /z/ endings which occur after vowels. (The /z/ endings after vowels and the /əz/ ending will be discussed later.)

Figure 5 shows that the Japanese learner is sensitive to two or possibly three more environmental constraints preceding the endings: [sonorant], [coronal], and possibly [continuant]. When we look at the first major division, [sonorant], we note that the percentage of target production for Subjects A and C is always higher after nonsonorants than after sonorants. That is, [-sonorant] facilitates the production of the correct /s/ or /z/ ending, while [+sonorant] hinders such production. This observation is consistent with the above finding for voiced and voiceless pairs. That is, since all [+sonorant] segments are also voiced, we would expect lower percentages than for the [-sonorant] category. The effect seen for Subjects A and C, however, does not hold for Subject B. Observe that Subject B's production in all categories, except /t/ and /d/, is much higher than the equivalent categories for Subjects A and C. Thus, it may be that after reaching a fairly high degree of accuracy, the [sonorant] constraint is no longer relevant. Or it may be a personal ideosyncrasy. This conjecture warrants further examination.



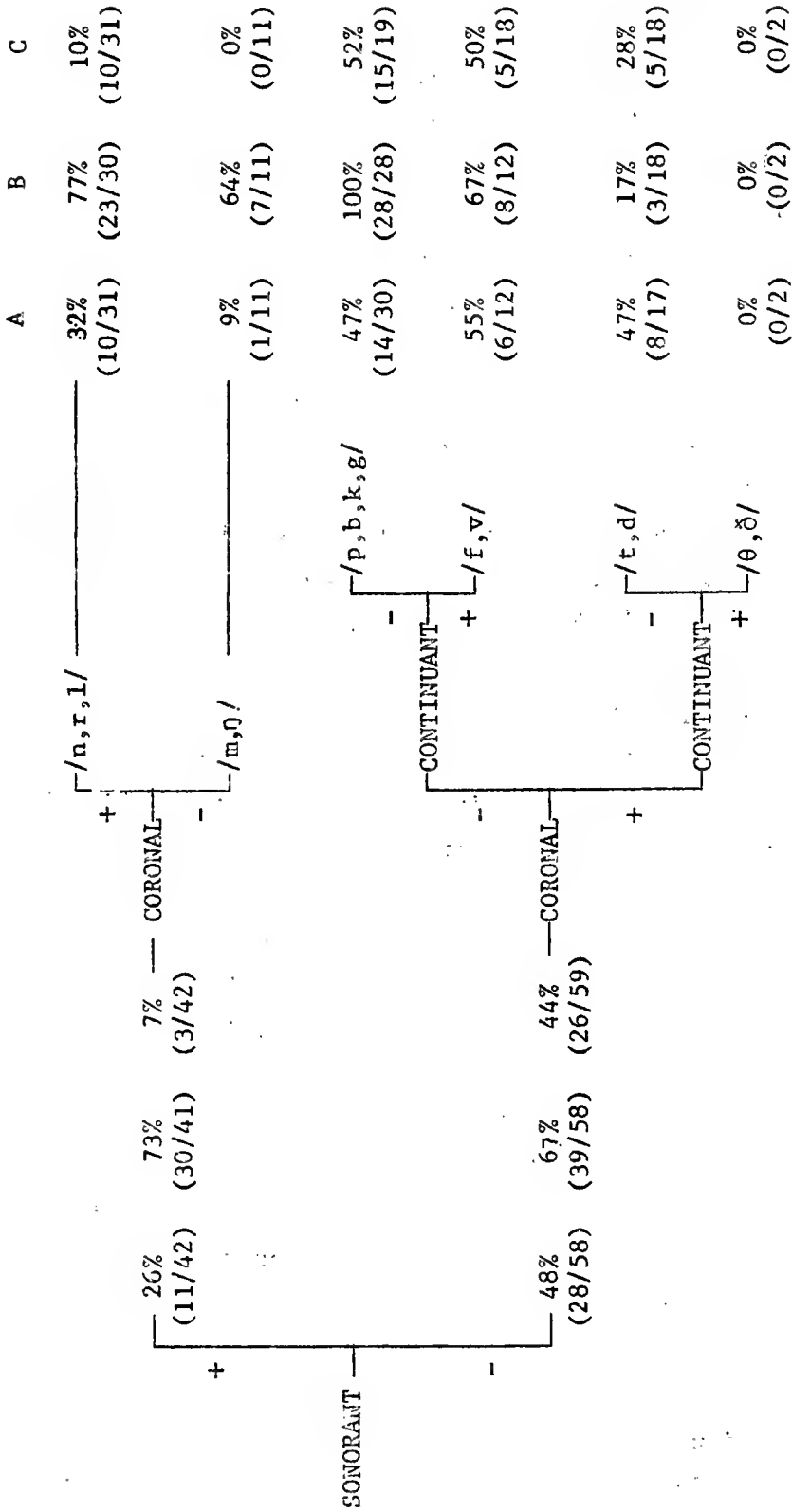


Fig. 5. Percentages of Correct Production of /s/ and /z/ Categorized by Preceding Environmental Constraints

The [voice] and [sonorant] distinctions, however, are not the only dimensions of sensitivity. Also highlighted in Figure 5 is a third dimension, [coronal]. Note that within sonorants, each of the three subjects has more target production after [+coronal] consonants (/n,r,l/) than after [-coronal] consonants (/m,ŋ/). Stated differently, within the category [+sonorant], target production is facilitated when the point of articulation of the final consonant of the noun is the same as the /s/ or /z/ ending, as in the alveolar sequences /nz/, /rz/, /lz/. Target production is hindered when the two points of articulation differ, as in the labial + alveolar sequence, /mz/, or the velar + alveolar sequence, /ŋz/.

The opposite effect is seen for [-sonorant] segments. The [+coronal] feature (/t,d/) constrains target production; the [-coronal] feature (/p,b,k,g,f,v/) encourages target production. Thus, in the [-sonorant] category, a difference in point of articulation between final stem sound and ending (e.g., /ps/, /gz/, /fs/, /vz/) promotes success; similarity in point of articulation (e.g., /ts/, /dz/) impedes success. This regularity holds for all three subjects.

The data for Subject B strongly suggests a fourth dimension of possible environmental sensitivity: [continuant]. Note that Subject B scores considerably higher when the final segment of the singular noun is [-continuant] than when it is [+continuant]. Compare 100% (/p,b,k,g/) with 67% (/f,v/). The fact that only Subject B (whose performance in four out of five categories is superior to the performance of the other subjects) shows a sensitivity to the [continuant] constraint may indicate that this dimension is more relevant for the Japanese speaker who is well on his way to mastering the target articulations for the three endings. This possibility requires further investigation.

From Figures 4 and 5 it is clear that a systematic organization does underlie the use of the /s/ and /z/ endings. The phonetic characteristics of the stem-final segment profoundly influence the learners' articulatory success with the two short {Z} allomorphs. For all three subjects, there is greater success when the final stem sound is voiceless than when it is voiced. For the [continuant] and [coronal] features, all subjects display remarkable patterning. For all subjects, [sonorant] is the primary con-

straint, and [coronal] is the secondary constraint. And for all, success is favored most when [sonorant] and [coronal] match in polarity, either ++ or --. Subjects A and C are alike in showing greatest target production when the stem consonant is [-sonorant], as demonstrated in Figure 6. By contrast, Subject B shows greatest target production when the stem consonant is [+sonorant], also demonstrated in Figure 6. The only difference, then, between Subject B and the other two subjects is a reversal in the polarity of [sonorant], which favors success.<sup>6</sup>

HIERARCHY OF CONSTRAINTS	SUBJECTS		HIERARCHY OF CONSTRAINTS	SUBJECT B
	A	C		
[-son, -cor]	48%	54%	[+son, +cor]	77%
[-son, +cor]	47%	28%	[+son, -cor]	64%
[+son, +cor]	32%	10%	[-son, -cor]	51%
[+son, -cor]	9%	0%	[-son, +cor]	10%

Fig. 6. Hierarchical Ordering of Constraints Affecting Target Production for Subjects A and C versus B

To complete the detailed study of all three endings, we will now look at pluralization of the final two categories of nouns, those ending in vowels and those ending in sibilants. Figure 7 below shows that the categories sibilant + /əz/ (in boxes) and vowel + /z/ (in circles) bear no consistent relationship to other categories from subject to subject. However, this does not mean that no patterning exists. For all three subjects there is one clearcut regularity--for /əz/ endings after sibilants there is more target production than for /z/ endings after vowels. For each subject, compare the percentages in boxes with the percentages in circles.

Stem-Final Consonant	SUBJECTS		
	A	B	C
/p,b,k,g/	68%		
	50%		
/f,v/	47%	100%	52%
		78%	
/t,d/	55%	67%	50%
		55%	32%
/n,r,l/	47%	17%	28%
			20%
/m,ŋ/	32%	77%	10%
	9%	64%	0%

Fig. 7. Comparison of Percentages for Sibilant + /əz/ (boxes) with Vowel + /z/ (circles)

In answer to Question Two, which deals with the influence of a phonological environment upon the specific ending, we have observed patterned regularities in the performance of each of the three subjects. Now, for a final comparison, let us contrast the findings from the two points of view--the two-way, long/short approach and the three-way, /əz/, /s/, /z/ approach. Figure 3 presents group performance for the data displayed earlier in Figure 5.<sup>7</sup>

Stem-Final Environment	I % Correct Long/Short Approach	II % Correct /əz/, /s/, /z/ Approach
[+son] {	/n, r, l/	39%
	/m, ŋ/	24%
[-son] {	/p, b, k, g/	66%
	/f, v/	57%
	/t, d/	30%

Fig. 8. Percentages Correct for Two-Way and Three-Way Distinctions

Based on Figure 8, three observations are in order: First, the percentages for Column I (two-way distinction) are higher in every case than the corresponding percentages for Column II (three-way distinction). The sharp differences in these figures seem to reflect two different learning tasks, each somewhat independent of the other. Specifically, it appears that one of the students' tasks is to learn the basic requirements of plural attachment (long/short). The other task is to learn the pronunciation rule for the phonetic rendering of the plural endings. The comparison in I and II suggests either that the learners have progressed faster with their first task than with their second, or that they find the first task easier than the second.

The second observation based on Figure 8 is that the two highest categories in Column I (both [+sonorant]) rank at or near the bottom in Column II. Clearly, it is more difficult for the learner to maintain like voicing for endings following [+sonorant] segments than for endings following [-sonorant] segments. It is not clear, however, why the [+sonorant] categories are the ones to show such wide fluctuation. However, the drastic reversal from very high to very low percentages for the [+sonorant] categories adds further support for the idea that there are

two separate learning tasks involved.

The third observation is that the [-sonorant] categories maintain the same ranks in relation to each other in Columns I and II. The most likely explanation of this fact is that the environmental constraints, such as [sonorant], [coronal], and [continuant] affect the learner's performance at both levels of processing, plural attachment and correct phonetic rendering of endings.

#### IMPLICATIONS FOR PEDAGOGY

The systematic behavior of learners as they use the plural morpheme can be of practical value to the ESL teacher in areas such as appreciation of the nature of language learning, evaluation of progress, and classroom presentation of materials.

The ESL teacher can gain from this study a more realistic view of the influences affecting the learner's performance. In particular, he/she should see that the use of noun plurals is not characterized by random success and failure, but rather by patterned performance. Furthermore, the learner's success with pluralization is neither categorically good nor categorically bad. Instead, it is more or less successful depending on the allomorph involved and its environment.

This view should affect the teacher's classroom expectations. Specifically, the teacher should not expect uniform performance from students for all three noun plural endings or for different environments of one ending. And when the teacher observes students who are struggling for mastery of these endings, the mistakes should not be looked upon as only a random hodgepodge; rather, the mistakes should be viewed as important pieces of an overall pattern (Dickerson and Dickerson 1978).

Since more than one learning task seems to be involved, the teacher should count it as progress when students are successful at attaching a long or short ending to a noun, even if *pages* is pronounced as /peɪjəs/ and *pies* is pronounced as /paɪs/. (In fact, because it is a more attainable goal and communication is hindered only slightly if at all, many teachers may want to consider correct long/short production as their only goal, rather than push their students for a more accurate pronunciation of

the three separate endings.)

The ESL teacher should also realize that for accurate evaluation of student progress, he/she must use a sufficient number of test items, and items which represent the full range of environments in which the noun plural is used. To test by using too few items or to test by not using items representative of each relevant category means that the learner cannot show all he knows, the teacher will not see all he knows, and the test results will be skewed.

Although most classroom teachers have some general ideas about which of the endings are easier or more difficult, there is now a stronger basis for sequencing teaching items along an easy to difficult continuum. However, since language background may be a factor influencing ease and difficulty, the teacher should not assume that endings which are easy (or difficult) for speakers of one language are easy (or difficult) for speakers of another language.

In short, informed by the findings of studies such as this one, classroom teachers should be better equipped for the basic tasks of materials development, classroom presentation, and testing.

#### SUMMARY

This study, dealing with the use of the noun plural morpheme {Z} by three Japanese speakers, has shown each subject to be operating from a language system which produces highly organized behavioral patterns. The fact that each subject finds long plurals easier than short plurals calls into question the claim that the short plural form is universally easier than the long. The fact that all three subjects show sensitivity to various environmental constraints highlights for the first time the role of phonetic environment in plural usage (and presumably also in plural acquisition). And the fact that the subjects' accuracy in attaching a plural ending differs so much from their accuracy in producing the correct pronunciation of those endings suggests that there may be two different learning tasks involved in the acquisition of noun plural endings. Findings such as these can be of value to the language acquisition researcher, the materials developer, and the classroom teacher.<sup>8</sup>

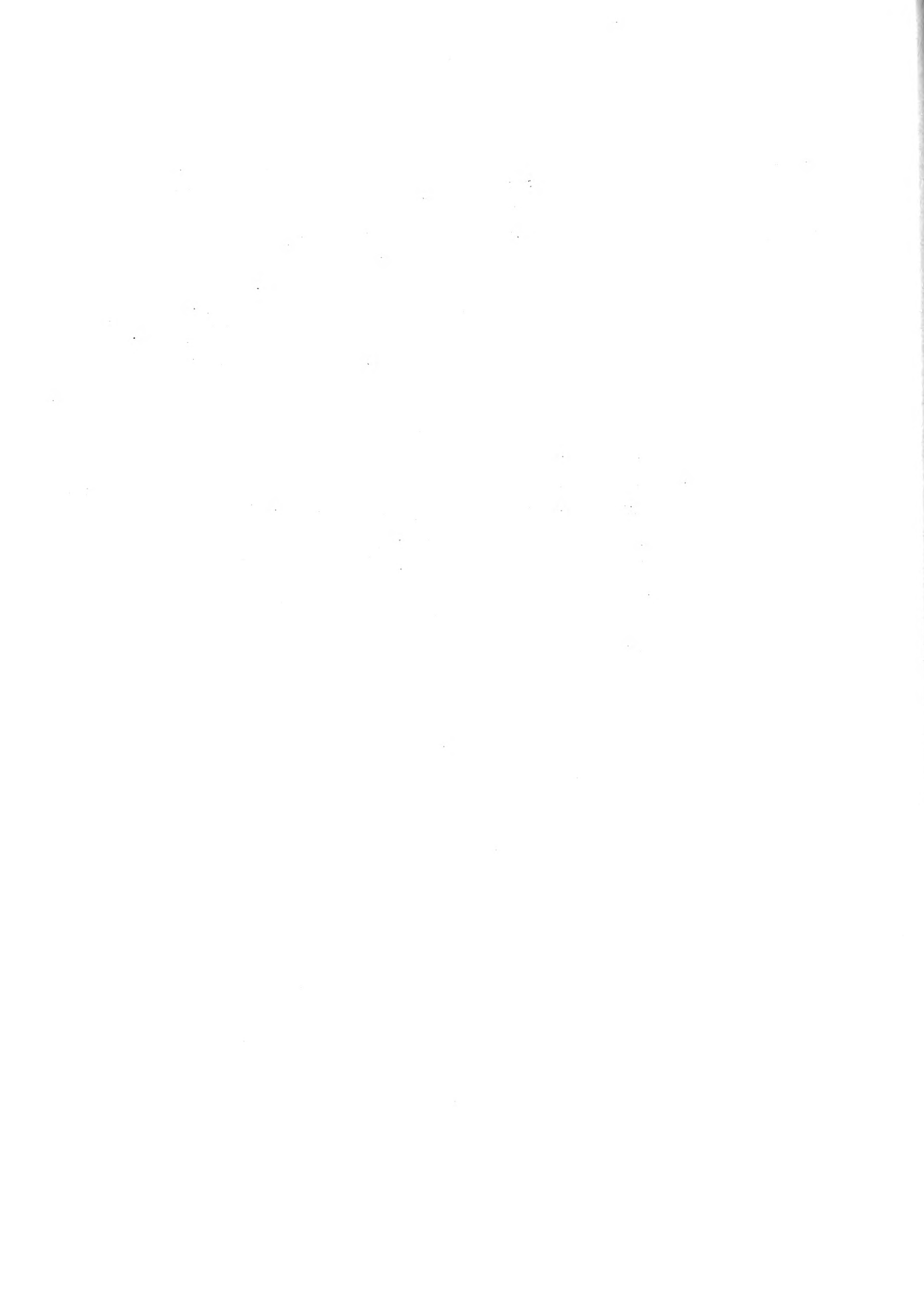
## FOOTNOTES

- <sup>1</sup>Subjects were also controlled for non-English foreign language training. Japanese speakers with extensive exposure to other foreign languages were not considered as potential subjects. The three subjects chosen for this study had so little exposure to other foreign languages that it was deemed insignificant in terms of its potential influence on their performance.
- <sup>2</sup>As will be shown later in this paper, depending on phonological environment, some plurals, both long and short, are easier to produce than others. When so few items are used, the researcher may inadvertently select a disproportionate number of either easy or difficult test items.
- <sup>3</sup>Since /ʒ/ is not obligatory in word-final position in English, the consonants /ʒ/ and /j/ were grouped into one category. (For example, words like *garage* and *corsage* can be pronounced with either /ʒ/ or /j/ as the final consonant sound.)  
 Final /ð/, as in *scythe*, was not tested, as its plural form is also variable in native speech, such that the final /ð/ often drops out entirely when the plural is formed.  
 Although items with word-final /θ/ were included in the test, these items were dropped from the analysis because the subjects did not produce /θ/ in the test items.
- <sup>4</sup>Incorrect responses were of two types: either no plural ending was used, or the long plural ending was attached to a noun which required the short plural.
- <sup>5</sup>A more detailed look at individual variation will appear later in this paper.
- <sup>6</sup>We are grateful to Wayne Dickerson for these observations about environmental constraints and the relative weighting of constraints for the three subjects. From such hierarchies, it is possible to write variable rules, in the manner of sociolinguistics research, to describe the learner's interlanguage phonology. For an example of such rule writing and a description of the model of research, see L. Dickerson (1975), L. Dickerson and W. Dickerson (1977), W. Dickerson (1976, 1977).
- <sup>7</sup>Since it closely parallels the group data, individual data is not presented here.
- <sup>8</sup>The authors are especially grateful to Wayne Dickerson for his criticisms of an earlier version of this paper.



## REFERENCES

- Bailey, N., C. Madden, and S. Krashen. 1974. Is there a "natural sequence" in adult second language learning? *Language Learning* 24.235-243.
- Dickerson, Lonna J. 1975. The learner's interlanguage as a system of variable rules. *TESOL Quarterly* 9.401-407.
- Dickerson, Lonna J. and Wayne B. Dickerson. 1977. Interlanguage phonology: current research and future directions. *The Notions of Simplification, Interlanguages and Pidgins and their Relation to Language Pedagogy*, ed. by S. P. Corder and E. Roulet. Geneva: Librairie Droz, 18-29. Also in *TESL Studies* 1977. 64-87.
- Dickerson, Lonna J. and Wayne B. Dickerson. 1978. Learning to pronounce: research responses to classroom concerns. *English Teaching Forum* 16:4.13-15.
- Dickerson, Wayne B. 1976. The psycholinguistic unity of language learning and language change. *Language Learning* 26.215-231.
- Dickerson, Wayne B. 1977. Language variation in applied linguistics. *ITL, A Review of Applied Linguistics* 35.43-66.
- Dulay, Heidi C. and Marina K. Burt. 1974. Natural sequences in child second language acquisition. *Language Learning* 24.37-53.
- Larsen-Freeman, Diane E. 1975. The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly* 9.409-419.
- Natalacio, D. C. and L. F. S. Natalacio. 1971. A comparative study of English pluralization by native and non-native English speakers. *Child Development* 42.1302-1306.



FIRST LANGUAGE PERCEPTUAL STRATEGIES AND READING IN A SECOND LANGUAGE

Richard Bruce Rickard

Previous reports have introduced syntactically determined perceptual strategies as one manifestation of psycholinguistic guessing. Accordingly, such strategies must be language specific and any transfer from the L<sub>1</sub> will affect reading in an L<sub>2</sub>. Using a CA approach, specific hypotheses were formulated to test this premise via the L<sub>2</sub> performance of native Chinese, Japanese, Persian, and Spanish speakers reading English. Korean subjects were included without a priori hypotheses. Source sentences containing specific patterns of subordination or negation were projected for intervals designed to force rapid reading. Subjects then indicated which of 2 alternatives was a correct paraphrase. The data was analyzed to determine performance differences between native language groups, and within groups, in relation to items and patterns. Results showed significant correlation with TOEFL scores and demonstrated reliability between NL groups. Significant differences were isolated between NL groups for relative difficulty of particular items, and evidence was found for performance differences between patterns for the Japanese. Several hypotheses found support, as did the concept of interference from L<sub>1</sub> perceptual strategies affecting psycholinguistic guessing while reading in an L<sub>2</sub>. Several related questions are raised as suggestions for further research.<sup>1</sup>

INTRODUCTION

In 1970 Kenneth Goodman published his discussion of psycholinguistic guessing in reading, i.e., that the mind, while efficiently reading, is always leaping ahead, attempting the unknown at an adventuresome pace, moving aggressively through uncharted passages by ". . . partial use of available minimal language cues selected \* \* \* on the basis of expectation." The implications of such a model have since been further developed by a number of educators and psycholinguists, particularly several involved in foreign language teaching and having interests as well in contrastive analysis and first language (L<sub>1</sub>) transfer and interference in second language (L<sub>2</sub>) learning.

Reporting on experiments with subjects reading their native language (NL), Fodor, Garrett, and Bever (1974) showed that in addition to semantic relationships, a reader relies extensively on syntactic expectations; e.g., a strategy is described which says that for native speakers of

English the first noun--verb--noun pattern constitutes an actor--action--receiver relationship and the main clause unless otherwise marked. They reported that aberrations from this pattern resulted in comprehension difficulties for subjects.

Extending these concepts across language boundaries, other researchers have variously reported on efforts to integrate Goodman's claims and Fodor, et al's findings with such concepts as interlanguage, approximative systems, and language transfer--models of L<sub>1</sub> and L<sub>2</sub> acquisition and L<sub>2</sub> performance discussed by Selinker (1972), Memser (1971), and Richards (1972), respectively. These concepts accommodate, to varying degrees, a consideration for structural differences between languages, i.e., contrastive analysis (CA)

Cowan (1975, 1977), in an effort to clarify relationships between CA and psycholinguistic guessing, hypothesizes that a successful reader has discovered, within his native language, principles for decoding meaning, such as those reported by Fodor, et al; and that furthermore, he has practiced skills which allow him to read in the idealized manner outlined by Goodman. However, because grammatically-based strategies are ipso facto language specific, a potential weakness is inherent when reading in an L<sub>2</sub>. Using examples of English contrasted with Japanese, Persian, and Hindi, Cowan demonstrates how reliance on L<sub>1</sub> strategies might affect the performance of someone learning to read in an L<sub>2</sub>.

Takayanagi (1975) refers directly to surface structure differences between Japanese and English which affect reading. Among those mentioned are the basic S-V-O English word order versus the common S-O-V Japanese order; differences in deletion rules for identical verbs; idiosyncrasies of relative clauses; and several problematic contrasts involving negation--one related to the position and range of the negative morpheme in each of the languages, and the other concerned with a Japanese speaker's ability to cope with negative subordinators, specifically 'unless'. Although Takayanagi makes no attempt to formalize a model and provides no substantive data with which to evaluate her claims, she asserts that these structural differences can be directly related to observable processing difficulties; and furthermore, that "radically different surface structures

may have some universal predictive value."

Brownscombe (1977) conducted a study examining the reading comprehension of native speakers of Japanese, Persian, and Spanish in relation to various types of relative clauses in English. She found evidence that native speakers of Persian, a language which invariably and explicitly marks relative clauses by inclusion of a pronominal referent within the clause to specify the antecedent, consistently have difficulty reading (i.e., correctly comprehending) reduced relative clauses in English--presumably because of the lack of explicitly designative morphemes to which they are accustomed.

### THE STUDY

It was with the intention of obtaining more substantive data about interfering L<sub>1</sub> perceptual strategies and of contrasting unique performance abilities of various NL groups reading English as an L<sub>2</sub> that the study reported herein was undertaken. Negation was selected as a framework for experimentation; it was chosen because it represents a linguistic universal (Greenburg, 1966) which, it was posited, would minimize the problems of equivalence associated with a CA approach (cf. Kachru, 1975). Also, as Klima (1964) has shown, negation in English is in no way simple, but rather has the capacity to manifest itself in a remarkable number of complex structural positions and morphemes, thereby suggesting practical reasons to evaluate learners' abilities to cope with it. Moreover, even a cursory glance at a few other grammars (Chao, 1968; Kuno, 1973; Lambton, 1953; Stockwell, et al, 1965; and Cho, 1975) shows that for Chinese, Japanese, Persian, Spanish, and Korean respectively, negation is an equally complex part of the language, and one which has at once both universal and idiosyncratic features. These considerations, as well as polarity and other properties (Baker, 1970) allowed negation to serve well as a medium for a number of interesting and testable hypotheses.

Subsequently, 6 specific patterns were targeted; they were:

- P-1 Negative subordinators and adverbials
- P-2 Conditional sentences with verbal negatives
- P-3 Negative determiners in subject (and actor) NPs

P-4 Negative determiners in object and complement NPs

P-5 Multiple negation within a clause

P-6 Multiple negation across clausal boundaries

In relation to these basic pattern groups, a number of predictive hypotheses were formulated in regard to the reading performance of native speakers of Chinese, Japanese, Persian (Farsi), and Spanish. The hypotheses were based on a priori CA, using written grammars, and were further refined through discussions with informants from each group. In part, these a priori hypotheses predicted that, among all 4 of the groups:

H<sub>1</sub> Double negatives expressing affirmative meaning (P-5 and P-6) would be the most difficult patterns of the 6.

Between groups, it was anticipated that:

H<sub>2</sub> Spanish speakers would perform nearest the norms of native speakers of English.

H<sub>3</sub> Double negatives expressing affirmative meaning would be particularly difficult for Japanese and Spanish speakers.

And considering the individual performance of each NL group relative to the 6 patterns, it was thought that:

H<sub>4</sub> Japanese speakers would have particular difficulty with the pattern 'negative determiner + NP' in subject and object position, and that this difficulty would be even greater when the pattern coincided with other negative morphemes to express an affirmative meaning.

H<sub>5</sub> Both Chinese and Japanese speakers would experience more difficulty comprehending a sentence where a negative subordinate clause followed the main clause than where it preceded the main clause.

#### METHOD

All test items were placed on 35mm. black-and-white slides, which

were then projected onto a screen via a remote-controlled projector. Based on preliminary studies, "source" sentences containing one of the target patterns were projected for 3 seconds--an interval determined to be sufficient for most subjects to read the sentence, but without enough time for extensive re-reading or reprocessing. (Control was maintained for sentence length: 15-17 words.) Immediately after the disappearance of the source sentence, subjects were shown a pair of sentences. They had previously been instructed to indicate the (1) sentence which was a correct paraphrase or expansion of the preceding source. The other sentence was always a distracter expressing a plausible misunderstanding of the source--usually an approximate antithesis. Subjects were allowed 12 seconds to read the pair. (Each of these alternatives contained only 10-12 words.) The screen then went blank and they had an additional 8 seconds to mark their response on a machine-scoreable answer sheet and to ready themselves for the next item. Following the presentation of every tenth set, the subjects were given a 1 minute rest break. The rigidly controlled timing was achieved by electronically pulsing the projector with a recorded program.

With the intended exception of their reading of the source sentences, the subjects seemed to function quite easily within the allotted time constraints. Since both of the secondary sentences--the correct answer and the distracter--largely recycled vocabulary from the original sentence, and since the observed behavior of the subjects indicated that they had more than enough time to process the antithetically paired alternatives, the assumption is that any error present should be attributed to faulty comprehension of the original source--faulty comprehension resulting from a failure to efficiently perceive and/or understand grammatical signals to meaning, i.e., inaccurate psycholinguistic guessing deriving from inappropriate L<sub>1</sub> perceptual strategies.

Following the planned design, 2 examples were presented, followed by 48 randomly ordered sets of test items. Of the 48, 39 items contained one of the target patterns; 9 were distracters containing no negative morphemes, randomly included. The subjects were 123 students enrolled in classes of the Intensive English Institute or the Division of English as a Second

Language of the University of Illinois at Urbana-Champaign. The research instrument was administered during the course of 3 consecutive days to individual classes meeting in their regular classroom under otherwise normal conditions. Subjects had no advance notification and were told only that the test was to evaluate and aid the development of the ESL curriculum, and that it would not affect their individual grade or evaluation. All classes received identical instructions and examples presented by the author. The teachers of each class played no active role in the administration, but were asked to remain present and to take the test themselves, for purposes of obtaining a control group of native speakers. The obtained sample included 16 Chinese, 9 Japanese, 23 Persians, and 16 Spanish speakers in addition to 19 Korean speakers and 40 speakers of a miscellany of less numerous languages. The unexpectedly large number of Koreans in the obtained sample resulted in the inclusion of that particular data into the study even though no predictive hypotheses had been developed in relation to that language. Although the sample cannot be considered homogeneous in English proficiency, some general relationships are tenable, and are further defined in the presentation of results and subsequent discussion.

The obtained responses were machine-scored; compiled output included the subjects' class placement level, sex, NL, and his/her total correct responses to the 39 target items (TESTTOT), as well as binary output for individual items--indicating right/wrong responses by subject, and by item for all subjects. To this was added the subjects' TOEFL total score (TOEFLTOT) along with subscores for reading (RDG), structure (STR), and vocabulary (VOC)--where such scores were available. All this data was then subjected to statistical analyses to probe relationships between and within NL groups in respect to items and patterns.

## RESULTS

Percentage means of TESTTOT ranged from 64.2% correct to 89.1% for the 5 NL groups now under analysis; means for the 39 individual items ranged from 48.1% to 98.8%. The mean for all subjects across all 39 items was 78.5% correct. These statistics indicate a reasonable level of



difficulty and reflected a normal distribution. There was also a significant positive correlation between the TESTTOT and the TOEFLTOT. The control group consisted of 7 native speakers of English; of these, 5 obtained perfect scores, while 2 erred on 1 item each. A summary of test results and correlative TOEFL data is shown in Table 1.

TABLE 1

Summary of research instrument results  
contrasted with available TOEFL statistics

Group	TESTTOT*				TOEFL**								
	%	$\bar{X}$	S.D.	n	$\bar{X}$ TOT	S.D.	n	VOC	n	RDG	n	STR	n
All subjects	78.5	30.6	5.7	123	470.9	66.5	79	40.7	45	47.7	54	44.9	54
5 NL groups	76.6	29.9	5.7	83	474.9	69.1	56	41.9	34	47.5	40	45.1	40
Chinese	89.1	34.8	3.1	16	536.9	42.8	9	58.0	4	57.8	4	54.3	4
Japanese	66.6	26.0	3.1	9	483.0	51.5	8	39.8	5	49.7	6	48.3	6
Korean	31.4	31.7	5.0	19	503.2	59.2	5	38.0	2	48.3	3	48.0	3
Persian	64.2	25.0	4.0	23	420.0	61.1	21	34.3	15	43.4	18	41.2	18
Spanish	81.9	31.9	5.7	16	505.1	40.3	13	50.4	8	49.2	9	45.8	9

\* 100% = 39

\*\* The correlation between TESTTOT and TOEFLTOT was .611 for all subjects; .665 for the 5 NL groups ( $p < .01$ )

#### DIFFERENCES BETWEEN NL GROUPS

Table 1 indicates substantial differences in English proficiency between the NL groups; one-tailed t-tests between each group and the remainder of the sample showed that: 1) the Chinese achieved significantly better results on both the TOEFLTOT and the TESTTOT; 2) for the TESTTOT alone, the Japanese and Persians were of the other extreme; and 3) only the Persians indicated significantly lower English proficiency in relation to all other subjects as measured by the TOEFLTOT ( $p < .001$  for all tests). Accordingly, subsequent discussion takes account of these relationships between NL groups.

In an effort to isolate and describe performance differences between groups and within groups in relation to particular sentences and structures, the following observations were made by: 1) examining the rank correlation coefficients for all permutations of paired groups; 2) contrasting the hierarchy of the 10 most difficult items (10 MDI) for the respective groups; and 3) isolating the items demonstrating significant difficulty for a particular NL group in relation to the performance of all other subjects, or in relation to NL groups of similar or greater L<sub>2</sub> proficiency.

The respective ranks of items--hierarchically arrayed in direct relation to score--showed significant correlation between all permutations of NL group pairs and provide more evidence for the reliability of the test instrument, showing that it obtained reasonably consistent results. Table 2 shows the correlation coefficients.

TABLE 2

Spearman rank correlation coefficients\* between NL groups

	Chinese	Japanese	Korean	Persian	Spanish
Chinese	---	.60	.56	.43	.67
Japanese		---	.64	.43	.72
Korean			---	.52	.81
Persian				---	.51
Spanish					---

\*All significant at  $p < .01$ .

With one exception (Korean--Spanish) however, these correlations are only moderate and allow for appreciable constituent differences between ranks. Some of these differences were isolated by examining the abbreviated ranks of the 10 most difficult items for each group, ranked from most difficult to least difficult. Not surprisingly, many items appeared repeatedly in the respective lists of 10 MDIs; Table 3 shows the frequency with which pairs of NL groups found particular items commonly difficult, but more interesting is that each NL group had among its 10

MDIs at least 1 item unique to it alone.

TABLE 3

The number of items commonly difficult between pairs of NL groups or (uniquely) difficult to a language group

(Determined by the comparison of ranks of 10 most difficult items)

	<u>Chinese</u>	<u>Japanese</u>	<u>Korean</u>	<u>Persian</u>	<u>Spanish</u>
Chinese	(3)	4	6	4	6
Japanese		(1)	5	6	6
Korean			(1)	4	7
Persian				(2)	5
Spanish					(1)

The individual performance of each NL group was considered from another approach as well, by using one-tailed t-tests to isolate significantly lower NL group means for individual items. With the notable exception of the Spanish speakers, each NL group found from 1 to 17 items uniquely more difficult than did the other subjects. These findings are shown in Table 4.

TABLE 4

The number of items significantly and uniquely difficult to an NL group (Determined by one-tailed t-tests of NL group means, by item, against the means of all other subjects;  $P < .05$ )

Significantly more difficult for:	<u>Chinese</u>	<u>Japanese</u>	<u>Korean</u>	<u>Persian</u>	<u>Spanish</u>
Number of items	2	5	3	17	0

A more interesting view of these differences between NL groups can be had by considering the following observations. In relation to both TESTTOT and TOEFLTOT, Spanish speakers and Korean speakers demonstrated similar abilities, i.e., no significant difference in means; they also showed high rank correlation (cf. Table 2) and it can be seen from Table 3

that they shared 7 items among their lists of 10 MDIs. However, analysis also showed that apart from these commonly troublesome items, they differed strikingly on 2 other items; furthermore, Korean speakers did significantly less well than the Persians on one item even though Persians comprised the lowest general ability. Conversely, the Koreans' performance on yet another item surpassed even that of Chinese speakers, the group showing highest proficiency. And perhaps most dramatic of all these phenomena, the single most difficult item for the Chinese, and one of only 2 on which they performed significantly less well than did the other subjects, was conversely of the next-to-the-easiest rank for the Persians and was the only item they found significantly easier than did other subjects. This item is shown in Example 1, which also provides a look at one of the test items. The item scores (% correct) are also shown, in alphabetical order by NL group (i.e., Chinese, Japanese, Korean, Persian, and Spanish).

Ex. 1. Since it is protected by no back-up system, the reliability of the main system is critical.

62.5 77.3 84.2 91.3 73.3

A: Because there is a back-up system, the reliability of the main system isn't critical.

B: The reliability of the main system is critical because there isn't a back-up system.

Contrasting the performance of Japanese speakers and Persian speakers, who can also be said to represent approximately equal levels of English proficiency, obtains similar disparities. Table 5 summarizes the more radical disparities between NL groups, and the relevant examples are shown below.

Ex. 2. Unless optional limits are specified by the user, the basic program computes only eight basic statistics.

3. A State Department report indicates that in 1971 no U.S. computer corporations sold hardware to China.

4. Only if the customers pay their bills within thirty days will there be no service charges.

5. Although the tax bill was opposed by no special-interest group, it did not pass easily.

6. John's mother will call Dr. James in the morning unless John's fever has fallen by then.
7. The mountain climbers could not have reached the icy peak without their newly-designed experimental boots.
8. Ted's physics professor said that he knew of no published research which had proven the theory.
9. Susan told no one about her brother's embarrassing car accident in Chicago the weekend before last.
10. Unless the Federal Reserve Board lowers current reserve ratios, banks will have to raise interest rates.
11. References indicate there was no one in the immediate family who could not speak three languages.

TABLE 5

Divergent performance results between groups of similar proficiency and performance results inversely related to L<sub>2</sub> proficiency

Example	<u>Performance results*</u>	<u>Test pattern</u>	<u>English proficiency** relationship</u>
1	Persian > Chinese	P-2	Chinese > Persian
2	Korean > Spanish	P-1	No difference
3	Spanish > Korean	P-3	No difference
4	Korean > Chinese	P-2	Chinese > Korean
5	Persian > Korean	P-6	Korean > Persian
6	Japanese > Persian	P-1	No difference
7	Japanese > Persian	P-5	No difference
8	Persian > Japanese	P-4	No difference
9	Persian > Japanese	P-4	No difference
10	Persian > Japanese	P-1	No difference
11	Persian > Japanese	P-6	No difference

\* (one tail;  $p < .05$ )

\*\* (Determined by t-tests of TOEFLTOT and/or TESTTOT; two tails,  $p < .05$ )

## DIFFERENCES WITHIN NL GROUPS

Examples 9 and 11 (above) and 12 and 13 (below) represent evidence in support of H<sub>4</sub>, which predicted that Japanese speakers would have particular problems with 'negative determiner + NP' in subject or object position, and that this problem would be magnified when this structure coincided with the presence of another negative morpheme to produce affirmative meaning. The scores for these examples are displayed in Table 6.

- Ex. 12. In times of national emergency, no idea is unworthy of consideration in order to protect human lives.
13. Legal experts say that no city will be unaffected by the new law passed last month.

TABLE 6

Relative scores for Japanese on representative items containing the pattern 'negative determiner + NP'

<u>Example</u>	<u>Chinese</u>	<u>Japanese</u>	<u>Korean</u>	<u>Persian</u>	<u>Spanish</u>	<u>Test pattern</u>
9	100	66.7	100	95.7	93.8	P-4
11	100	66.7	89.5	95.7	87.5	P-6
12	93.8	33.3	78.9	39.1	80.0	P-5
13	87.5	44.4	77.8	40.9	62.5	P-5

Table 6 is also representative of the Spanish speakers and multiple negation (cf. H<sub>3</sub>)--while they exhibited occasional difficulty with patterns 5 and 6, it was neither consistent nor significant. (Similarly, between groups, the Spanish speakers demonstrated neither consistently nor significantly lower scores on these two series of items.) And finally, H<sub>5</sub> (Chinese and Japanese: negation and alternate orders of conditional clauses) found no support.

## DISCUSSION

The most obvious success was proving the methodology itself: the experimental technique was effective in controlling extraneous variables

and in eliciting easily isolated performance differences capable of being subjected to intensive a posteriori analysis. Moreover, the isolation of so many significant performance differences between NL groups--particularly between more-or-less equally proficient groups and between inversely proficient groups--provides considerable support for the concepts of psycholinguistic guessing and interfering perceptual strategies shaped by L<sub>1</sub> factors.

Furthermore, several predictions did find support. Indeed, for all but the Chinese, pattern 5 (multiple negatives within one clause) did prove the most difficult group of sentences (cf. H<sub>1</sub>). Another hypothesis which found support--indirectly--was H<sub>2</sub>: that Spanish speakers would perform nearest the norms of native English speakers because of the unique familial relationship between the 2 languages; although they did not achieve the highest mean score of the 5 groups and were not of the highest proficiency level, it is to be recalled that Spanish speakers found no item significantly more difficult than did other subjects, and that they were the only group for which that is true (cf. Tables 1 and 4). The most strongly substantiated hypothesis however, was H<sub>4</sub>--relating to Japanese speakers, negative NPs, and multiple negation expressing affirmative meaning.

It must not be overlooked that many variables remain untested and many questions are still unanswered: e.g., What is the relationship between perceptual strategies, comprehension, and time? Can comprehension be improved at the price of time-consuming reprocessing? Also, it must be remembered that sentences are not usually read in unrelated isolation, nor can comprehension be categorically tied to sentence level grammar. To reflect the "real" reading situation, perceptual strategies should also be explored in relation to longer rhetorical units, e.g., sequential sentences, paragraphs, short passages, etc. Only by considering such potentially relevant factors can evidence for psycholinguistic guessing and perceptual strategies be developed with validity. And assuming that additional evidence for these concepts is found, still more questions remain: What in the L<sub>1</sub> figures most prominently in the development of perceptual strategies? What, if any, are the universal characteristics? How much is transferred to the L<sub>2</sub>, and what to the learner's advantage,

what to his disadvantage? And more easily answered perhaps, will be questions examining relationships between interfering strategies and L<sub>2</sub> proficiency e.g., are inappropriate strategies abandoned as L<sub>2</sub> proficiency increases, or do they constitute an ineradicable fossilization? These questions and many more must be empirically explored and answered if psycholinguistic guessing and perceptual strategies are ever to be understood and substantiated.

\* \* \* \* \*

#### FOOTNOTES

- <sup>1</sup>The research reported herein was conducted while the author was a Teaching Assistant in the Division of English as a Second Language of the University of Illinois at Urbana-Champaign. The author wishes to acknowledge the valuable advice of J Ronayne Cowan during the development of this project and to extend thanks for his criticism of an earlier version of this paper. An abridged version of this paper appears in *TESOL Quarterly* 13.4 599-602.



## REFERENCES

- Baker, C. L. 1970. Double negatives. *Linguistic Inquiry* 1.169-186.
- Brownscombe, E. Carol. 1977. Toward characterizing perceptual strategies of second language learners: a pilot study. In Y. Kachru (ed.), *TESL Studies* 2. 20-50. Urbana-Champaign: Division of English as a Second Language, The University of Illinois.
- Chao, Yune-Ren. 1968. *A Grammar of Spoken Chinese*. Berkley: The University of California Press.
- Cho, Choon-Cho. 1975. The scope of negation in Korean. In Ho-min Sohn (ed.), *The Korean Language: its structure and social projection*, 48-62. Honolulu: The Center for Korean Studies, The University of Hawaii.
- Cowan, J R. 1975. Reading, perceptual strategies, and contrastive analysis. In Y. Kachru (ed.), *TESL Studies* 1. 24-37. Urbana-Champaign: Division of English as a Second Language, The University of Illinois.
- Fodor, J. A., T. G. Bever, and M. F. Garrett. 1974. *The Psychology of Language*. New York: McGraw-Hill.
- Goodman, Kenneth S. 1970. Reading: a psycholinguistic guessing game. In H. Singer and R. B. Ruddell (eds.), *Theoretical Models and Processes of Reading*, 259-271. Newark, Delaware: International Reading Association.
- Greenburg, Joseph H. 1966. *Language Universals*. The Hague: Mouton and Company.
- Kachru, Yamuna. 1975. Toward defining the notion 'equivalence' in contrastive analysis. In Y. Kachru (ed.), *TESL Studies* 1. 82-98. Urbana-Champaign: Division of English as a Second Language, The University of Illinois.
- Klima, Edward S. 1964. Negation in English. In J. A. Fodor and J. J. Katz (eds.), *The Structure of Language*, 246-323. Englewood Cliffs, New Jersey: Prentice Hall.
- Kuno, Susumu. 1973. *The Structure of the Japanese Language*. Cambridge, Massachusetts: The MIT Press.
- Lambton, Ann K. S. 1953. *Persian Grammar*. Cambridge: University Press.
- Nemser, William. 1971. Approximative systems of foreign language learners. *International Review of Applied Linguistics* 9.115-123
- Richards, Jack C. 1972. Social factors, interlanguage, and language learning. *Language Learning* 22.159-188.
- Selinker, Larry. 1972. Interlanguage. *International Review of Applied Linguistics*. 10. 219-231.
- Stockwell, Robert P., et al. 1965. *The Grammatical Structures of English and Spanish*. Chicago: The University of Chicago Press.

Takayanagi, Fumie. 1975. Structural differences in Japanese and English which affect learning to read. In Y. Kachru (ed.), *TESL Studies* 1. 115-121. Urbana-Champaign: Division of English as a Second Language, The University of Illinois.

REVIEWS

PRINCIPLES OF LANGUAGE LEARNING AND TEACHING. H. Douglas Brown.  
Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1980. Pp. xi + 276.

Brown's Principles of Language Learning and Teaching (PLLT) is not a book to be dipped into for information on a particular topic related to the learning or teaching of foreign languages. Read in that way, it yields mostly summaries of and references to key books and articles. The effect is then disappointingly superficial; but that is the reader's fault, not the author's. Nor is PLLT the usual cookbook text for language teachers. It says nothing about technique and very little about method. It is, instead, a unique and timely discussion of what language learning is and how language teachers can approach their task. It interweaves knowledge and viewpoints from the fields of education, psychology, and (not even primarily) linguistics, and it deserves to be read from cover to cover as the author intended.

Within its 276 pages the reader, who is presumed to be a student of language pedagogy, encounters arguments and counterarguments about the main issues surrounding language learning and teaching as they have evolved over the past several decades, principally but not exclusively in the United States. Brown clearly favors the viewpoints of cognitive psychology and generative semantics as overriding principles for language teaching and learning activities, yet he never preaches or talks down to the reader about them. In fact, despite the frequent use of personal pronouns to establish a relationship between author and reader, I would suspect that most student teachers will find it difficult to feel that comfortable with some parts of the book because of its comprehensiveness, its distillation and synthesis of significant theories from many sources. Yet, with the help of classroom discussion and some of the readings suggested at the end of each chapter, the serious student should come away with an enviable grasp of the nature of language and the process of learning it, i.e., the theoretical foundations of foreign language teaching upon which to base a variety of classroom techniques with some confidence.

The book is organized into twelve chapters of about twenty pages each with the following titles: Language, Learning, and Teaching; First Language Acquisition; Comparing and Contrasting First and Second Language Acquisition; Human Learning; Cognitive Variations in Language Learning; Personality and Language Learning; Sociocultural Variables; Comparing and Contrasting Two Languages; Error Analysis--The Study of Learners' Interlanguage; Discourse Analysis--The Study of the Pragmatic Functions of Language; Foundations of Measurement and Research; and From Theory to Practice. At the end of each chapter is a briefly annotated list of suggested readings, most of which have been touched on in that chapter, and from seven to twelve topics and questions for study and discussion which focus on the new information and relate it to issues presented in earlier chapters. (One typographical error on page 63 refers the reader to a quote on the wrong page--47 instead of 57. There are several other typos, but most are inconsequential.)

Besides a thirteen-page bibliography, there are two helpful indexes: index of names and index of subjects.

Cutting across the chapter topics and recurring frequently throughout the book in various guises are eight major issues, as identified in the author's preface:

- competence vs. performance
- innateness (the nature-nurture controversy)
- universals vs. variability
- language and thought
- imitation (surface vs. deep structures)
- the role of practice
- input and feedback

By addressing these issues from several theoretical standpoints, the author has an opportunity to support or refute various hypotheses from the works of Whorf and Sapir, Skinner, Fries, Chomsky, Piaget, Carl Rogers, and others, and thus to show steady development in the foreign language profession's understanding of the complexity of language and of human learning. In my opinion, no other book for students of foreign language teaching takes such a broad view of these complex issues and yet treats

them so clearly and fairly as this one.

The result is a text whose plan is "to build toward a comprehensive, integrated understanding of the teaching-learning process, such that one will be able to construct a personalized rationale, or theory, of second language acquisition" (p. viii)--surely a unique approach among textbooks for educating foreign language teachers today.

Thus, the author's stated purpose is neither to suggest how to teach a foreign language (although he outlines the rationales behind several methodologies in the final chapter) nor to advocate one theoretical stance above all others. Rather, what Brown offers here is a kind of guided tour of foreign language territory, leading the reader from one theoretical milepost to the next toward a fuller understanding of the language learner's task and the teacher's responsibility, including the role of classroom research in furthering that understanding for the individual teacher.

Some critics might complain about the book's lack of practical guidelines for the novice and its assumption that everyone who teaches should understand a broad range of issues in related disciplines. Yet I find this book refreshing and stimulating for just those reasons. As Brown says:

The language teaching profession has been through its stage of childhood and adolescence, its grasping at a method here and a technique there, and its wild claims of final, clear-cut answers. It now stands in young adulthood, secure in its general understanding of the language-acquisition process... but...prepared to face the complexity and unpredictability of human behavior... (p. 245).

That coming of age and this text's unique approach make PLLT an ideal book for the profession to begin the 1980's with. A thorough grasp of its content will give us all a sharper perspective on where we have been and what paths to knowledge our professional inquiry might follow in the near future.

Barbara F. Matthies



LIVING LANGUAGE: USA CULTURE CAPSULES FOR ESL STUDENTS. Jerilou Johnson. Rowley, Mass.: Newbury House Publishers, 1979, Pp. 115.

Living Language: USA Culture Capsules for ESL Students fills a long standing void in the ESL field, that of lack of an English text designed for high intermediate and advanced levels which attempts to acquaint the student with various aspects of United States culture while developing his language skills. Optimally this book is to be used in the ESL classroom within the United States, but it may also highly benefit those students, business personnel, and other persons in an EFL environment who wish to "brush up" on their use of English and their understanding of American culture in anticipation of an extensive trip to the United States.

The premise upon which Living Language is based is that the advanced student of English coming to this country for the first time generally has two major problems. The first is that he is unpleasantly shocked to discover that his English simply isn't working here for reasons he's not quite sure of. The second "problem" is that he most likely arrives here with a whole artillery of pre-conceived and often well formulated conceptions about the United States which will either be accurately borne out in the days to come or he will be "shocked" to discover that even these pre-conceptions don't work, don't seem to fit. The frustration of this kind of culture shock may well force him to abandon the "truth" of American life as he had conceived of it; in short, it may lead him inside himself even to a questioning of his own basic values.

This then is where Living Language begins. In her preface to the student, Johnson states that many of the problems as described above are "...unnecessary, and the frustrations you experience trying to resolve them use an enormous amount of energy which could be directed to more productive goals. An awareness of some basic life styles in the United States can enable students learning English to feel more comfortable as they try to function in a new culture." (vii). She goes on to express the belief that after completing these ten units, or culture capsules, the student's "ability to communicate will not be limited by (his) lack of cultural understanding of life in the United States".

The approach in Living Language is thus one primarily based on the idea that a lesson must be meaningful to the student to be well learned. Each one of the ten capsules is an extended dialogue, usually between two people, centering around one theme. The ten themes, or unit titles, are as follows: 1) City Streets, 2) Woman Alone, 3) A Father's Day Gift, 4) Eating Right, 5) Living in a Loft, 6) Staying Fit, 7) The Blacker the Berry, 8) Blue Collar Blues, 9) Apartment Hunting, 10) Family Clean-Up. Within these ten dialogues, a plethora of topics are introduced and discussed, including unemployment, single's life, shopping, education, urban development, divorce, racism and discrimination, family life, and the role of women.

Though it remains first and foremost a book on culture, a fair number of new vocabulary words, phrases, and even slang are introduced and, at the end of each chapter, drilled. These vocabulary words are to be learned in context. The student is not asked to be able to spell and define each word, but rather is asked to be able to recognize its meaning in context. That is, vocabulary work is an outgrowth of the dialogues and situations presented.

Johnson suggests that the dialogues be listened to on tape by the student before the vocabulary drills and comprehension questions at the end of each chapter are attempted. She does not specify whether such tapes are available; most likely it will be up to the teacher to make her own. The author stresses that the student should not be expected to read the dialogue or take part in role-playing of it before hearing it read either by the teacher or by a native speaker on the tape. Any kind of lifeless, mechanical reading has no place here: the author seeks not only the promotion of cultural awareness and introduction of new vocabulary in these capsules but also the development of the oral production and listening comprehension skills of expression, pronunciation, and intonation. Intonation, in fact, plays a vital role in the interpretation of various degrees of subtle humor and sarcasm present in several of the dialogues. For example, capsule one describes an event which took place on the first day of a young Mexican student's arrival in Chicago. His difficulty in finding a taxi to take him into the city, and his



subsequent dialogue with the driver end with the student's dismay and astonishment at the fare he must pay:

Driver: ...Well, here we are. That'll be \$14.30, tip not included.

David: \$14.30!!!!!!!?????????

Driver: Man, that's why I said you should have taken the limo. It only costs three dollars and it comes right to this hotel. Maybe that's why your friend told you to come to this hotel.

David: Maybe. Here's \$15.00. Keep the change.

Driver: Look, man, it's okay this time...you're just a student, and a stranger here and everything. But just so you know in the future--seventy cents ain't no tip on a \$14.30 fare!

Only on a well-prepared tape or upon a well-prepared reading could the correct intonation and phrasing signal David's concern and the driver's astonishment. Students thus learn to extract meaning from suprasegmental aspects of communication as well as from the written word.

At the end of each dialogue eight to ten comprehension questions are sequenced according to events or situations described within the dialogue. These are concerned with the student's inference and understanding of the actions and speech acts which took place rather than asking for a mechanical recall of details. If, after hearing the dialogue one or several times, the student can answer the comprehension questions without referring back to the dialogue, he should have a good idea of the main ideas and arguments at hand.

The vocabulary review section at the end of each chapter reintroduces the new vocabulary of the dialogues into contexts other than those of the dialogue. The student is asked to select from the list of words that which appropriately fills the blank in the sentence. At the end of every three capsules, a noncumulative vocabulary review section tests the student's recall of the new vocabulary presented in those chapters with the same fill-in-the-blank format.

Johnson presents the vocabulary solely for the sake of the meaning and coherence of the dialogues. She thus advises that "the majority of time and teaching effort should be directed to Questions for Discussion and Suggested Activities" (xii). The teacher brings her own resources to

Questions for Discussion. Living Language would lose much of its effectiveness taught by a non-native speaker. Yet Johnson does provide for this loss by including an Instructor's Notes section which provides additional comments and clarifications of the capsules. These notes also provide advice to the teacher on guiding the discussion for the individual capsules and they predict problems and misunderstandings which may creep up.

Discussion and free exchange of ideas on controversial issues is the focal point of Living Language. In the discussion questions students are often asked to compare facets of American culture with their own culture and in doing so to identify and examine points of similarity or difference between the two (a cultural error analysis?). Thus they compare the culture which they know with the culture they want to learn about, and all by using their second language, English. For an example of the types of discussions arising in this book, the following are from the capsule entitled "Eating Right":

-Compare shopping in a U.S. supermarket with the way people normally shop in your native country. Are there any supermarkets in your native country? If so, compare them with the ones you have seen in the U.S. or the one described in the dialogue.

-Judy is obviously interested in the current trend toward natural foods. What do you know about this? Do you eat any natural or diet foods? Which ones? Why?

-People in the U.S. have been severely criticized for extensive waste. Do you think this criticism is justified? Give reasons and examples. Is there waste in your native country?

Suggested Activities following capsule nine, entitled "Apartment Hunting" include the following:

-Discuss the differences in the living style of the young people in the U.S. and the young people in your native country.

-Have you ever tried to rent your own apartment? Share your experiences with other students.

-Bring to class pictures from magazines which show the kind of kitchen, dining room, living room, bedroom, etc. that you'd like to have in your "dream apartment". Describe them to the class.

Many of these activities involve trips into the community and are thus better used in an ESL environment in the USA. However, by way of adaption, the teacher can choose to skip any one of the Suggested Activity questions that may not be appropriate. The teacher can also skip any one of the capsules: they are not sequenced according to content although some of the vocabulary words learned and stressed in one chapter may re-enter unemphasized in a subsequent chapter. Also if the instructor is using the dialogues for role play in the classroom and there is an all-male or all-female dialogue, he may elect to omit the capsule or to adapt it to a more objective third person perspective. In essence, then, Living Language is a supplementary text in that it would be better used as only a fraction of an integrative class that also works with other texts in the areas of grammar, reading, pronunciation, etc. Perhaps, given a typical one-semester five hour a week class, one culture capsule a week on an appointed day with approximately an hour and a half to two hours of in-class time and a comparable amount of time spent in student preparation prior to the class should be sufficient to complete the book without dwelling too long on one subject. Such a once a week affair could provide a rich and rewarding break or diversion for the students and may well become the highlight of their week, for a well-handled discussion of the nature intended by this book would touch upon the affective side of the individual student, allowing him to experience concentrating more on meaning than form in English in expressing his own opinions, frustrations, and questions. This appropriate loss of the unconscious monitor system most of the advanced English students carry around with them will allow them to communicate more freely in their second language.

Activity suggestions at the end of each chapter lend themselves well to full-scale term projects and the teacher may want to encourage students to, at some point in the semester, pick a topic which particularly interests them and to pursue this topic in more depth. Such projects can send a student into the community and lead to real interaction with native speakers. The only caution in assigning such long term projects is that the teacher should make sure that a student is interested in the undertaking and thus is highly motivated. Assigning assignments at random

would seem to defeat the whole theory of meaningful learning inherent in the format of this book. Since a student's motivation for performing a task directly affects his performance on the task, a teacher should allow a student ample leeway in his selection, including helping him to design a project of his own.

Living Language contains a balanced variety of speech styles. These registers are vitally important, especially to an advanced level student for whom few listeners will make allowances for social blunders if that student already has a good command of the language. In attempting to present actual cultural situations of American life Living Language also does a nice job in providing realistic dialogues in these situations, and in doing so employs many different functions of language. The following excerpts illustrate this point:

Blue Collar Blues (Capsule 8)

Waitress: Yeah, what'll it be?  
 Willard: Cup of coffee, Honey.  
 Waitress: Here you are. Anything else?  
 Willard: I'll have a chocolate-covered doughnut, too.  
 Waitress: Right!

(Later, she returns)

Waitress: Fill it up?  
 Hank: Yep, go ahead.

(And later)

Willard: Can we have the checks, Honey?  
 Waitress: Here you are...forty-five cents.  
 Willard: There you go...  
 Waitress: Thanks. Have a nice day.

Staying Fit (Capsule 6)

Nurse: Good afternoon. May I help you?  
 Shirley: Yes, I have an appointment to see Dr. Smyth at 3:30.  
 Nurse: Your name, please?  
 Shirley: Shirley Copeland.  
 Nurse: Oh yes, Ms. Copeland. Come right this way to room #2. Would you please put on this examination gown. The doctor will be with you in one moment.

Family Clean-Up (Capsule 10)

- Ms. Scott: Mary, John, Adam. Time to get up. Breakfast will be ready in ten minutes...Rise and shine.
- Mr. Scott: Do you really expect any response at 8:05 on a Saturday morning?  
(He walks to the stairs and calls)  
Breakfast will be on the table in three minutes.  
The last person to be seated will wash the breakfast dishes.
- (There is a loud noise and the thunder of feet is heard as the three Scott children race downstairs to the breakfast table.)
- Adam: I'm first, Dad. I got here first. No dishes for me.
- Mary: ....I'm second.
- John: You are not second. You got to the table second, but I sat down before you did.
- Mary: Daddy said, "To the table", he didn't say "To sit down".
- John: He said, "To be seated." Didn't you, Dad?

Woman Alone (Capsule 2)

- Mary: So, the boys are adjusting O.K. How about you?
- Grace: You know, Mary, I feel relieved and actually excited. The life of the young divorcee is not a bowl of cherries, but I feel like a whole person for the first time since I can remember. Certainly there are more pressures; running a home alone is not easy. But recently I've been painting much more...
- Mary: That's good to hear, Grace. Are you seeing anyone, anyone special, that is?
- Grace: Oh, there are several interesting men in my life; but I don't feel that frantic need to "find somebody". I'm really quite content; the boys, my job at the university, my painting...The next time--if there is a next time--I don't think there will be this desperate need to depend on someone..

Blue Collar Blues represents a realistic waitress in contrast to the very formal "May I take your order?" phrases usually found in notional-functional type instructional materials. Particularly of note in this dialogue are the endearment terms loosely used by the truckers which serve an interactional function here of ensuring social maintenance. They would definitely not serve this function and therefore would not be appropriate in other situations. Short, clipped phrases are also

prominent in this dialogue, from the "there/here you are/go" phrases to the "fill it up?" "yep" exchanges.

Staying Fit is an example of a formal, consultative style and includes many interactional speech acts in a more formal vein than those above. Upon greeting Shirley, the doctor (who is a woman) asks her "How are you?", to which Shirley answers "Fine", and then proceeds to explain her health problems and the reason for the visit.

Family Clean-Up is in an intimate style and aptly portrays an example of children's speech as well as logic. It also includes some often used phrases like "rise and shine", "time to get up", "I got here first".

Woman Alone is again in an intimate style, but this time the dialogue is between two close friends discussing a very personal, emotional subject: divorce. Here we see Mary in the role of a sympathizer and also a questioner; we see her probing gently for information that Grace needs to be coaxed out of. In this conversation as well as the others the student hears several other important conversational features like topic nomination, topic development, and finally topic termination. Although the important non-verbal dimensions of discourse cannot be assessed from the written or taped dialogue, they can be brought out in role playing in the classroom. Another way of using the dialogue in the classroom is to make inferences, on the sole basis of the conversation alone, as to what the personality of the participants are. Students will often see a character in a completely different light than the teacher does and by discussing these opinions, the teacher can gain insight into helping a student cope with and adjust to the new culture.

Living Language has as its primary goal the teaching, or rather, the presenting, of American culture to foreign students. While culture is a broad term covering all the ideas, customs, skills, and arts shared by a group of people, the author of this book has chosen to delimit the concept to the area of, specifically, American values, and to those values which seem to be the most controversial. One issue that is dealt with several times is the changing role of women in the U.S. In chapter two we see a newly divorced woman holding down a job, working on her Ph.D. and trying to keep her family together at the same time. In chapter four two house-

wives meet in the supermarket and consequently the talk runs from food to prices to children to Girl Scout meetings. Other chapters deal with various other life styles--two young students struggling to survive on a tight budget, a young wealthy couple remodeling a loft, a black couple concerned that they are being discriminated against by the white mechanic fixing their car, etc. The author seems to have accomplished her goal of not making value judgements on United States culture, but of simply presenting small capsules of that culture and encouraging the student to think for himself, and perhaps to realize that there are basic ideas and attitudes toward life that are totally different than those of his mother country. The fact that many types of relationships between people, such as woman-woman, man-man, student-student, children-parents, husband-wife, etc., and many different age groups and income brackets are exemplified allows every learner from young adolescent through adult business executives to find a point of identification to begin with and with which to compare and contrast his own country. By thus encouraging an open mind this book may, if used by a sensitive teacher, lead into some really productive discussions and may allow a student to feel free to air his own "culture shock" frustrations, anger, and estrangement which in many cases are a result of misunderstandings about the target culture. Hearing the experiences of his peers may also help him realize that the feeling of anomie, of homelessness, often experienced at his level of language learning is a very real feeling and one that can be accepted and dealt with in positive ways.

Within the dialogues, different registers and styles of speaking are represented, and this "slice of life" technique is, in my opinion, very natural and cohesive, accurately reflecting the status and education of the participants. I found that not one of the ten dialogues seemed forced or simply strung together in order to get a cultural point across. Each dialogue is interesting because the points discussed are so controversial and changing. The title page of each capsule is illustrated with a black-and-white drawing depicting the people in the situation described. In addition, other "visual aids" including an example of a charge account application, a student health form, a job application, an unemployment

compensation form, and an apartment lease accompany appropriate chapters and are useful in that they are forms a foreign student will most likely run into at some point in his stay in the U.S.

Finally, regarding the secondary function of this book to introduce new vocabulary words, I think that the inclusion, of "slang" expressions and words is most appropriate and well-handled. I personally believe that higher level students should be exposed to slang words and realize that they are not necessarily "bad". In fact, in some registers they are appropriate. The key phrase here is "some registers" it is these registers that must be pointed out to the student. I think Living Language makes a good beginning attempt at this much overlooked area of English teaching and learning. Some of the slang terms introduced in the various dialogues include, "wow!", "cracking down", "yep", "it sure beats...", "in a jiffy", "catch you later", "what's up?", "get off my back", and others.

In summary, I find little to negatively criticize about Living Language. I have found that there is a dearth of advanced material available in the ESL field, especially advanced material that is designed specifically to help students overcome the culture shock they all experience in varying degrees. This book teaches a foreign student what he wants to know, what's really important for him to know as a human being trying to exist in a world quite different from his own. The teacher-student interaction encouraged by the discussion provides the teacher with valuable feedback on the student's needs and wishes, and allows both students and teachers to develop the interpersonal relationship Carl Rogers feels is so vital to learning of any type.

Mary R. Reichardt





# *TESL STUDIES, 1980*

*LYLE F. BACHMAN*

*The Construct Validation of Oral Proficiency Tests*

*H. DOUGLAS BROWN*

*The Role of Teacher Feedback in Prevention of The Fossilized Errors of Second Language Learners*

*GARY A. CZIKO*

*Psychometric and Edumetric Approaches to Language Testing: Implications and Applications*

*WAYNE B. DICKERSON*

*A Pedagogical Interpretation of Generative Phonology I: Theoretical Foundations*

*CYNTHIA A. GOCKLEY*

*Toward Developing an Optimal Test of Writing Proficiency*

*BRAJ B. KACHRU*

*Models for New Englishes*

*YAMUNA KACHRU*

*'Transfer' in 'Overgeneralization': Contrastive Linguistics Revisited*

*SUSAN OSUCH-HATZIAVRAMIDIS  
LONNA J. DICKERSON*

*Phonological and Morphological Conditioning of {Z} in Speech of Japanese ESL Students*

*RICHARD B. RICKARD*

*First Language Perceptual Strategies and Reading in a Second Language*

## *REVIEWS*

*BARBARA MATTHIES*

*H. DOUGLAS BROWN: Principles of Language Learning and Teaching*

*MARY REICHARDT*

*JERRILOU JOHNSON: Living Language: USA Culture Capsules for ESL Students*

1157

# *TESL*

## *STUDIES 1981*



4

PUBLICATION OF THE DIVISION OF ENGLISH AS A SECOND LANGUAGE  
UNIVERSITY OF ILLINOIS

# *TESL STUDIES*

*EDITORS: YAMUNA KACHRU, J RONAYNE COWAN*

*EDITORIAL BOARD: Katherine O. Aston, Lyle F. Bachman,  
Lawrence F. Bouton, H. Douglas Brown, Lonna J. Dickerson,  
Wayne B. Dickerson and Rebecca G. Dixon*

*TESL STUDIES is intended as a forum for the presentation, in pre-publication form, of the research done by the faculty and students of the Division of English as a Second Language, University of Illinois, Urbana-Champaign. TESL STUDIES will also publish research of other University of Illinois faculty and students if it is of relevance to our field. All interested in obtaining copies of TESL STUDIES should write to the following enclosing payment (Checks should be made payable to the University of Illinois): Mrs. Norma H. Robinson, Division of English as a Second Language, 3070 Foreign Languages Building, Urbana, Ill. 61801.*

*PUBLICATION OF THE  
DIVISION OF ENGLISH AS A SECOND LANGUAGE  
UNIVERSITY OF ILLINOIS  
URBANA, ILLINOIS*

*Price per copy: \$4.00*

# ***TESL STUDIES***

## ***EDITORS***

***YAMUNA KACHRU  
J RONAYNE COWAN***

## ***EDITORIAL BOARD***

***Katherine O. Aston; Lyle F. Bachman  
Lawrence F. Bouton; H. Douglas Brown  
Lonna J. Dickerson; Wayne B. Dickerson  
Rebecca G. Dixon***

***VOLUME 4  
1981***

***DIVISION OF ENGLISH AS A SECOND LANGUAGE  
UNIVERSITY OF ILLINOIS  
URBANA-CHAMPAIGN***



TABLE OF CONTENTS

		Page
Lawrence F. Bouton	The Imperative Tag: What is It Really Like?	1
Carol Chapelle Joan Jamieson	ESL Spelling Errors	29
Lonna J. Dickerson	Evaluating, Selecting and Adapting Pronunciation Text-books: I. The Evaluation and Selection Process	37
Wayne B. Dickerson	A Pedagogical Interpretation of Generative Phonology: II. The Main Word Stress Rules in English	57
Stephen B. Dunbar	Models of Human Memory in Second Language Learning	95
Joan E. Friedenberg Curtis H. Bradley	Micro-ESL: A Method for Teaching the Functional/Notional Syllabus and Developing Communicative Competence	115
Stephen M. Smith	Acting Methods Applied to the Teaching of English as a Second Language	127
<i>REVIEWS</i>		
Fred G. Davidson	RONALD MACKAY, Bruce Barkman, and R. R. Jordan, eds. <i>Reading in a Second Language: Hypotheses, Organization and Practice</i>	141
Mary Siekert	JOAN MORLEY: <i>Improving Spoken English</i>	147





TABLE OF CONTENTS

		Page
Lawrence F. Bouton	The Imperative Tag: What is It Really Like?	1
Carol Chapelle Joan Jamieson	ESL Spelling Errors	29
Lonna J. Dickerson	Evaluating, Selecting and Adapting Pronunciation Textbooks: I. The Evaluation and Selection Process	37
Wayne B. Dickerson	A Pedagogical Interpretation of Generative Phonology: II. The Main Word Stress Rules in English	57
Stephen B. Dunbar	Models of Human Memory in Second Language Learning	95
Joan E. Friedenberg Curtis H. Bradley	Micro-ESL: A Method for Teaching the Functional/Notional Syllabus and Developing Communicative Competence	115
Stephen M. Smith	Acting Methods Applied to the Teaching of English as a Second Language	127
<i>REVIEWS</i>		
Fred G. Davidson	RONALD MACKAY, Bruce Barkman, and R. R. Jordan, eds. <i>Reading in a Second Language: Hypotheses, Organization and Practice</i>	141
Mary Siekert	JOAN MORLEY: <i>Improving Spoken English</i>	147



THE IMPERATIVE TAG:  
WHAT IS IT REALLY LIKE?

Lawrence F. Bouton

An effective derivation of the tag appended to imperative stems to form sentences like Sit down, will you? has proved elusive. A copying rule of the sort that produces the tag in He's here, isn't he? on the basis of the statement form in the stem cannot be used to generate imperative tags; it cannot account consistently for the latter's subject, auxiliary, or polarity. Nor do other proposed derivations such as the WHIMPERATIVE-based approach of Sadock (1970) prove more effective.

In an attempt to find solutions to the problems raised by these different analyses of the imperative-plus-tag construction, we turn to such factors as the intonation of the imperative tag and the polarity of the imperative stem as a means of predicting characteristics of the imperative tag that have up to now seemed unsystematic. In the process, we will find that if we are to take advantage of this new predictability of the imperative tag, we must generate that tag at the initial level of the derivation of the imperative-plus-tag construction. Furthermore, we find that an effective approach to the syntax of this construction must take into consideration syntactic, semantic, pragmatic, and perhaps phonological, characteristics of the imperative stem and its tag.

Finally, having described the imperative-plus-tag construction and suggested a derivation that seems to resolve many of the problems untouched by earlier analyses, we go on to discuss briefly the relevance of these findings to the ESL teacher.

INTRODUCTION

In English, the imperative is sometimes followed by a tag construction consisting of an auxiliary verb and its subject. The effect of this tag is frequently to soften the "imperativeness" of the stem, giving it some of the characteristics of a question and making it more polite than the bare imperative would be. Attempts have been made to link this imperative-plus-tag sequence transformationally to other structures in English, most notably, to the regular tag question (Katz and Postal, 1964), and to WHIMPERATIVES (Sadock, 1970). But these attempts, which will be discussed later, have failed in each case because the authors failed to consider certain characteristics of the imperative-plus-tag that their analyses later proved unable to account for. The purpose of this paper is to look

once more at the imperative-plus-tag sequence, with special attention being given to its possible subjects and auxiliaries, its polarity, and the intonation patterns that it can carry. In doing this, we will discover that we cannot deal effectively with the form and meaning of the tag if we treat it as being derived from a single source. What rules and constraints the tag seems to obey depends upon the polarity of the stem to which it is attached and the intonation contour of the tag itself. Very few characteristics are held in common by all tags on imperatives. Furthermore, there are characteristics of this tag in various contexts that must be handled by derivations that are to at least some extent ad hoc. What this means for the ESL teacher is that while some helpful transfer of learning can be expected to occur between the tag in the tag question and that in the imperative-plus-tag sequence, many of the features of the latter must be taught as idiomatic items to be memorized rather than derived through the use of general rules. It also means that since we will discover the imperative-plus-tag sequence to represent different constructions with different rules and constraints, the teacher must be careful not to confuse the students by inadvertently teaching these constructions together as if the sequence had a single unambiguous source.

#### PREVIOUS ANALYSES

Perhaps the first thing one notices about the tag attached to imperatives (1) is that it looks very much like that used to form tag questions (2). Both consist of an auxiliary followed by a subject; both are sometimes positive, sometimes negative.

- (1) a. Give me a hand here, won't you?  
b. Toss me that pen, will you?
- (2) a. You'll give us a hand when it's time, won't you?  
b. She has left already, hasn't she?  
c. Sally won't still be waiting for us, will she?  
d. They dropped out of the running, didn't they?

On the basis of the apparent similarity between the tags in (1) and (2), it is tempting to try to derive both by essentially the same rule. As a start, we might take note of a fact that is generally recognized, namely, that the tag used in conjunction with a statement to turn it into a question is quite regular. The subject of the tag is the pronominal form of the subject of the main clause (henceforth, the stem); the auxiliary of the tag is the same as that of the stem, except when the stem has none, in which case the

auxiliary do occurs in the tag; and the polarity of the tag (i.e., whether it is positive or negative), is the opposite of that of the stem. The form of the stem of a tag-question, then, permits one to predict the form of the tag associated with it. If one assumed that the tag on imperatives could be predicted in the same way, then it would be necessary to assign imperatives like (1) an underlying structure in which both a subject and an auxiliary were present. It was this assumption on which Katz and Postal (1964) based their analysis of the imperative-plus-tag construction. Further, since the only tags that they found grammatical in subjectless imperatives contained the auxiliary will and the subject you, they assumed that it was this auxiliary and this subject that were to be found in the deep structure of such sentences. Imperatives like (1), then, were to be derived from a structure something like (3-a), from which the tag formation rule would produce (3-b) and deletion rules (3-c).

- (3) a. \*IMP You will give me a hand.<sup>1</sup>  
b. \*IMP You will give me a hand, won't you?  
c. Give me a hand, won't you?

Through their analysis, Katz and Postal make the claim that the tag on an imperative is derived by basically the same rule as that in tag-questions, and that the tag on the imperative is, in this sense, quite regular. But further investigations into the imperative-plus-tag sequence have brought to light characteristics that make the Katz and Postal analysis untenable and show clearly that the tags on imperatives and those in tag-questions are not governed by the same rules. For example, Sadock (1970) pointed out what the reader may already have noticed: that in sentences like (1-b), the polarity of the tag is not the opposite of that of the stem, as would be true in a tag-question. Furthermore, Sadock went on, in such sentences, the polarity seems unpredictable, in that it may be the opposite of its stem, or it may not (4).<sup>2</sup>

- (4) a. Go home, will you?  
b. Go home, won't you?

Another difference between the tag on the imperative and that in the tag-question was brought out by Thorne (1966) and Dowling (1969) through examples like those in (5).

- (5) a. Everyone go home now, will you?  
b. Someone give me a hand here, will you?

It seems that the subject of the stem and the subject of the tag on an imperative do not always agree. The imperative stem may have overt subjects that are third person indefinite pronouns, but the subject of the tag in such sentences will invariably be you.

Still a third problem facing the Katz and Postal analysis was raised by Bolinger (1967) and Sadock (1970), when they showed that other auxiliaries than will can appear in the tag of the subjectless imperative:

- (6) a. Tell me what time he is leaving, can you?
- b. Help me lift this, would you?
- c. Stop by the office this afternoon, could you?

On the basis of examples like these, Sadock went on to argue that by deriving the tag attached to an imperative through the application of the same tag-formation rule used in deriving tag-questions, Katz and Postal would "predict erroneously that all imperatives are ambiguous between the sense of can and will, as well as between whether they contain the morpheme. . . PAST."<sup>3</sup>

From the discussion so far, we have seen that two of the three predictable characteristics of tag-questions that the tag-formation rule was set up to represent do not seem to be present in the tag associated with the imperative: the subject of the imperative tag does not seem to agree with the subject of the main clause in every case; and the polarity of the tag attached to an imperative stem will not necessarily be opposite to that of the stem itself. As for the auxiliary verb, as Sadock noted, if we derive it on the basis of the auxiliary of the imperative main clause, we must apparently assume that every subjectless imperative is ambiguous with regard to whether it has can or will as its auxiliary and whether its tense is present or past in its initial structure. For these three reasons, it seems apparent that attempting to generate the tags on the imperative and those forming tag-questions by the same rule is to claim a degree of similarity between the two tags that simply does not exist in English.

The failure of the Katz and Postal attempt to link the tag on imperatives to that in tag-questions led Sadock (1970) to propose an entirely different derivation for the former. He noticed that for every instance of what seemed to be an imperative-plus-tag that he considered, there was a counterpart in the form of a question, but with the function of an imperative (7). These counterparts Sadock labeled WH-imperatives, or WHIMPERATIVES.

- (7) Stop wiggling, will you? = Will you stop wiggling?

Sadock goes on to point out that these WHIMPERATIVES, like the imperative-plus-tag structures, have definite characteristics in common with simple imperatives. For instance, all three can co-occur with a following please (8) and with indefinite vocatives (9). None of the three can have a stative verb as its main verb (1). And all three can become infinitive complements of tell or ask in indirect speech.

- (8) a. Stand up, please.
  - b. Stand up, will you, please?
  - c. Will you stand up, please?
  - (9) a. Move that box, someone.
  - b. Move that box, will you, someone?
  - c. Will you move that box, someone?
  - (10) a. \*Grow tall.
  - b. \*Grow tall, will you?
  - c. \*Will you (please) grow tall?
  - (11) a. Open the door.
  - b. Open the door, will you?
  - c. Will you (please) open the door?<sup>3)</sup> }
- + He asked me to open the door.

But having shown that WHIMPERATIVES are imperative-like in several ways, Sadock reminded us that they were not really imperatives at all, but questions, and went on to argue that imperative-plus-tag sequences should be thought of in this way also. Using the line he perceived between WHIMPERATIVES and the imperative-plus-tag, Sadock managed to provide a derivation for the imperative-plus-tag that remedied some of the problems he found in the analysis by Katz and Postal.

The imperative-plus-tag construction, Sadock postulated, has the WHIMPERATIVE as an underlying structure. Given any particular WHIMPERATIVE, its imperative-plus-tag counterpart was formed by breaking the WHIMPERATIVE in two between its subject and whatever immediately followed it, and reversing the order of the two resulting segments (12). On the basis of this analysis, Sadock refers to sentences like (12-c), not as imperative-plus-tags, but as fractured WHIMPERATIVES.

- (12) a. Will you pick up some beer on the way home?
- b. Will you // pick up some beer on the way home?
- c. Pick up some beer on the way home, will you?

Sadock felt that his WHIMPERATIVE analysis of the imperative-plus-tag had remedied two of three weaknesses in the Katz and Postal thesis.<sup>4</sup> First, since both constructions seemed to permit or prohibit the same auxiliaries in sentences like those in (13), he could now account for the auxiliary

selections in both sentences by applying whatever constraints were involved only once, to the deep structure of the WHIMPERATIVE. Furthermore, each auxiliary in the tag has its own source in the underlying WHIMPERATIVE.

- (13) a. Will you answer the phone, please? = Answer the phone, will you?  
b. Won't you give me a break? = Give me a break, won't you?  
c. Could you let me know when he leaves? = Let me know when he leaves, could you?  
d. \*Must you stand there gawking, please? = Stand there gawking, please, must you?  
e. \*Shouldn't you go home, please? = Please go home, shouldn't you?

The second thing that Sadock's analysis let us do that the Katz and Postal analysis did not is to derive the positive and negative imperative tags each from its own underlying WHIMPERATIVE. In the process, he showed that the failure of the alternation between positive and negative tags on imperatives (14) to produce any correlated alternation in the polarity of the request itself is not totally ad hoc. The same thing is true of the WHIMPERATIVES (15). In neither case does the addition of the negative to the sentence change the essence of the request being made. In the (a) versions of (14) and (15), the effect of the negative is make the request more polite. In the (b) the negative makes the request appropriate in a situation in which the person asked is apparently hesitant or unwilling to do what he is being asked to do. What effect the change in polarity does have applies equally in the imperative-plus-tag and in the WHIMPERATIVE.

- (14) a. Play me some chess tonight, will you? = Play me some chess tonight, won't you?  
b. Hand me that wrench, can you? = Hand me that wrench, can't you?  
(15) a. Will you play me some chess tonight? = Won't you play me some chess tonight?  
b. Can you hand me that wrench? = Can't you hand me that wrench?

Sadock's derivation of the imperative-plus-tag from WHIMPERATIVES does account, then, for the construction as a whole is a question, at least in form. It also generates each of the tags containing a different modal from its own unique deep structure (its underlying WHIMPERATIVE), thereby avoiding the ambiguity found in this regard in Katz and Postal's analysis. And it avoids the dilemma posed by the alternation between positive and negative tags in the same environment by deriving the one by "fracturing" a positive WHIMPERATIVE, the other, by "fracturing" a negative. Further, by relating



the alternation between the tags on the imperative to that between positive and negative WHIMPERATIVES, and showing that the alternation produces the same variations in politeness in each of these sentence types, he makes the imperative tag less ad hoc. But there are several facts that Sadock does not attempt to explain. For instance, if one starts with a WHIMPERATIVE, there seems to be no way to "fracture" it in such a way as to produce sentences in which the auxiliary do appears in the stem as a carrier of the negative or of emphasis, e.g., (16) and (17). The WHIMPERATIVE counterpart of (16), for example, would be (18).

- (16) a. Don't look at me that way, will you?<sup>5</sup>  
b. Don't forget to mail my letter, will you?
- (17) a. Do try to be on time this time, will you?  
b. Do at least mention my name to your hiring agent, can't you?
- (18) a. Will you // NOT look at me that way?  
b. Will you // NOT forget to mail my letter?

Even if the do in (16) and (17) is part of "an invariable introductory formula" (Quirk, 1972, 7.73), or "a fossilized main verb lying outside the verb phrase" (Bollinger, 1967), and could be derived in some way so as not to interfere with the fracturing process, other problems remain in the path of Sadock's analysis. For instance, unlike negative WHIMPERATIVES, which permit the auxiliary to be either positive or negative (19), negative imperatives can take only positive auxiliaries in their tags (20). As a result, Sadock's analysis would incorrectly predict that (21-b) could be transformed into (20-b)

- (19) a. Will you // NOT take so long to finish this time, please?  
b. Won't you // NOT be so hard to get along with, for a change?
- (20) a. Don't take so long to finish this time, will you, please?  
b. "Don't be so hard to get along with this time, won't you?"

Another major weakness in the derivation of the imperative-plus-tag from WHIMPERATIVES is its inability to generate sentences like those mentioned earlier, in which there is a subject in the main clause as well as in the tag itself (21). There is no way that the fracturing rule can produce these sentences. This inability becomes even more obvious with sentences in which the main clause subject is a third-person indefinite pronoun and the tag subject is you (22).

- (21) a. You behave yourself, will you?  
b. Don't you say that again, will you?

- (22) a. Someone ask him how much it costs to fly home, will you?
- b. Everyone be quiet as you leave, won't you?
- c. No one say a word, will you?
- d. Someone figure out what we owe, can you?

Finally, we return to the claim made by Sadock concerning the essential nature of the imperative-plus-tag construction. As we noted earlier, he argues that the imperative-plus-tag is not an imperative at all; rather, derived from a WHIMPERATIVE, it is a question used imperatively. His strongest evidence for this claim involves the capacity of the various structures to co-occur with the speech-tag "I tell you." Sadock points out that while simple imperatives can take the tag (23), WHIMPERATIVES cannot (24). He then offers examples like (25) to show that the imperative-plus-tag behaves like a WHIMPERATIVE in this regard.

- (23)       Get out of here, I tell you!
- (24)       \*Will you get out of here, I tell you!
- (25)       \*Get out of here, will you, I tell you!

This evidence does seem to leave little doubt that the construction as a whole in (25) behaves like a WHIMPERATIVE. But the crucial question is not whether the entire construction is a question, but whether the stem itself is an imperative (as Katz and Postal thought), or the second part of a fractured WHIMPERATIVE. The answer lies in sentences like (26).

- (26)       Get out of here, I tell you, will you?

Here we find the stem itself quite able to co-occur with the speech-tag and behaving, in this respect, the way the simple imperative (22) behaves. The tag, will you, has been appended in its normal sentence-final position, making the overall sentence a question in some sense, but not affecting the essential nature of the stem itself. It appears that the sentences that we have been labeling imperative-plus-tag are just that, not fractured WHIMPERATIVES, after all. Sadock seems to be unable to account for the imperative-plus-tag construction, just as Katz and Postal were.

#### A NEW TACK

We seem to be back where we started. We are faced with a stem-plus-tag construction that looks somewhat like a tag question. A number of specific examples of these constructions have features that could be accounted for by using tag-question-like rules, as Katz and Postal suggested. Yet

there are a solid set of exceptions to such a rule that are also perfectly grammatical sequences. Nor are these exceptions handled effectively by the WHIMPERATIVE analysis. The problem is that none of the analyses done so far have asked whether there might be different identifiable contexts in which the imperative tag behaves in predictably different ways. If such contexts could be discovered, then the derivation of the tag in each of these contexts would be different, and what seem at this point to be arbitrary features of the imperative tag would be recognized as quite systematic.

Our next step, then, should be to see whether such contexts exist, and as we pursue this matter, we will pay particular attention to the intonation contour of the tag and the polarity of the stem. All the while, of course, we will be interested in developing rules that will generate the subject, the auxiliary, and the polarity of the imperative tag, and that will express the relationships that may exist between the imperative stem and its tag. We will come to see that the imperative-plus-tag cannot be approached as a single construction, derived by the same set of rules. Rather, different sets of imperative-plus-tag sequences, defined in terms of tag intonation and stem polarity, have different meanings, involve different constraints, and must have at least partially different derivations. As the initial source for these derivations, we will propose an initial structure in which the imperative tag is present, rather than trying to derive the tag by either copying or "fracturing" rules, as the previous analyses have done.

If we are going to explore the use of stem polarity and tag intonation contour as factors that can be used to define contexts to which rules generating imperative tags are sensitive, we should first describe what form these factors can take. Polarity is a straightforward concept: the stem may be either positive or negative. The possible intonation contours are somewhat similar to those of tag-questions, in that the tag itself may rise or fall in pitch, with a different meaning being assigned to each of the possibilities. The falling contour (31) is essentially the same for both the imperative tag and the tag of the tag-question, and it carries much the same meaning--a request for confirmation of what has just been said. In (27), the speaker indicates that he expects the addressee to agree that the statement in the stem is true; in (28), he indicates that he expects the addressee to agree to comply with what the stem asks of him.

(27) a. You're tired, aren't you?<sup>31</sup>

- (27) b. Not everyone is here yet, are they?<sup>31</sup>  
(28) a. Help Tim out whenever you can, won't you?<sup>31</sup>  
b. Don't let anyone get in here from now on, will you?<sup>31</sup>

When one tries to define the rising contours that can appear on imperative tags, on the other hand, he finds a much more complex pattern. The reader can develop a sense of this complexity very easily. Taking any imperative-plus-tag sequence in which the stem is positive, e.g., (29), pronounce the sentence with various slightly different rising contours on the tag. The reader will notice that at times those tags seem to reinforce the imperative tone of the utterance; at other times, to soften it.

- (29) a. Stand, will you?  
b. Stop by the office tomorrow, won't you?  
c. Give me a hand with this thing, would you?  
d. Pick out a new desk set for me sometime, can you?

Most of the contours that were uttered by the reader as he repeated the tags in these items fell into the general category defined by what can be labeled 3+. But whether the effect of this 3+ tag on the particular imperative is to increase or to decrease its demanding tone, there is a definite contrast between the impact of the 3+ tag and the tag with a falling contour--31. The latter expresses a confidence in the willingness of the addressee to comply with the request that is simply not found in the former.

This 3+ tag contour is not limited to tags on positive stems, however; for many native speakers, it can be appended to negative stems as well.<sup>6</sup> When this happens, there is a different contrast between the meanings signalled by the 3+ contour and that signalled by the 31. On a tag appended to a negative stem, the 3+ contour indicates that the speaker is referring to some action that is going on at the moment, and that he wants the addressee to stop doing whatever it is (30). The 3+ tag cannot be used on a negative stem if that stem is telling the addressee not to start something at some time in the future (31). In these latter sentences, the 31 pattern is appropriate.

- (30) a. Don't talk, will you?<sup>3+</sup> (= Stop talking!)  
b. Don't bounce that ball, will you?<sup>3+</sup> (= Stop bouncing that ball!)  
c. Don't keep dropping names, will you?<sup>3+</sup> (= Stop dropping names!)  
(31) a. Don't tell Pete about this mess we got ourselves into, will you?<sup>31</sup> (\*3+)

- (31) b. Don't forget to bring money for the tickets, will you?<sup>31</sup> (\*3+)  
c. Don't serve any port tonight, will you?<sup>31</sup> (\*3+)

There is one other intonation contour associated with the imperative tag: the 2+ pattern. Evidence of the difference between the 3+ and the 2+ patterns is found in the fact that sentences carrying the former cannot be paraphrased by regular tag questions used imperatively,<sup>7</sup> while the latter can.

- (32) a. You won't talk, will you? (= Don't talk, will you?<sup>2+</sup> (\*3+))  
b. You won't bounce that ball, will you? (= Don't bounce that ball, will you?<sup>2+</sup> (\*3+))  
c. You won't keep dropping names, will you? (= Don't keep dropping names, will you?<sup>2+</sup> (\*3+))

It is probably significant that the same 2+ tag contour found on the imperative stems in (32) is also found on the tag in the tag-question in that same set of examples. In having essentially the same intonation pattern as that found in the tag question, the 2+ pattern on imperative stems is like its 31 counterpart. They are also alike in that they both are used with imperatives indicating that the addressee is not to start something in the future, rather than that he is to stop something what he is doing. And finally, both the 2+ and the 31 patterns carry the same implication of certainty or uncertainty with regard to the addressee's response that they would carry in regular tag-questions. The 31 contour, we have already seen, signals that the speaker expects the addressee to comply willingly; the 2+ pattern suggests doubt, without carrying the imperious tone that the 3+ pattern does on occasion.

We have seen that there are three intonation patterns (2+, 3+, and 31) which can combine with two stem polarity patterns, giving us a total of six possible contexts defined by these elements. We will now take a look at how these contexts can be used to predict most of the variations in the form of the subject, the auxiliary, and the polarity of the imperative tag, as well as under what circumstances the imperative-plus-tag can co-occur with please, and when there can be both a tag and an explicit subject associated with an imperative stem. First, we will consider the tag subject, what form it can take, and how it is related to the stem subject.

The subject most frequently used in imperative tags is you, even when the explicit subject of the stem is an indefinite pronoun like someone,

\* everyone, or anyone (33).

- (33) a. Everyone go home now, will you?
- b. Someone shut the door, will you?
- c. Don't anyone say anything about this, will you?

So long as there is an explicit subject in the stem, nothing other than you can occur as the imperative tag subject. This invariability of the tag subject in this particular environment may be related to the fact that both it and its antecedent, the stem subject, refer to the addressee, but it cannot be explained by that fact. For example, when the subject of the stem is indefinite, other pronouns within the stem referring to that subject as their antecedent can take any of three forms: him/her, them, or you, or their variations.<sup>3</sup>

- (34) a. Someone cut  $\left\{ \begin{array}{l} \underline{\text{himself}} \\ \underline{\text{themselves}} \\ \underline{\text{yourself}} \end{array} \right.$  a piece of cake and see if it is any good.
- b. Everyone put  $\left\{ \begin{array}{l} \underline{\text{his}} \\ \underline{\text{their}} \\ \underline{\text{your}} \end{array} \right.$  books in this box.

One would expect the subject of a tag attached to one of these sentences to follow the same agreement rules as the other pronouns referring to the same antecedent, but such is not the case.

- (35) Someone cut  $\left\{ \begin{array}{l} \underline{\text{himself}} \\ \underline{\text{themselves}} \\ \underline{\text{yourself}} \end{array} \right.$  a piece of cake, will  $\left\{ \begin{array}{l} * \underline{\text{he?}} \\ * \underline{\text{they?}} \\ \underline{\text{you?}} \end{array} \right.$

Nor is this idiosyncrasy of the imperative tag subject typical of tag subjects in general. In tag-questions, for example, the tag subjects referring to indefinite antecedents do follow the same agreement rules as other pronouns in (36), even when the tag subject and its antecedent refer to the addressee, as it does in (36).

- (36) a. Someone bought  $\left\{ \begin{array}{l} \underline{\text{himself}} \\ \underline{\text{themselves}} \end{array} \right.$  a new car, didn't  $\left\{ \begin{array}{l} \underline{\text{he?}} \\ \underline{\text{they?}} \end{array} \right.$
- b. No one here had  $\left\{ \begin{array}{l} \underline{\text{his}} \\ \underline{\text{their}} \\ \underline{\text{your}} \end{array} \right.$  radio on all night, did  $\left\{ \begin{array}{l} \underline{\text{he?}} \\ \underline{\text{they?}} \\ \underline{\text{you?}} \end{array} \right.$

In (29)-(36), the underlined pronouns are all co-referential and all initially indefinite. When definitization applies to the second and third pronouns in each sentence on the basis of their antecedents, the form that the pronoun can take may agree with different features of the antecedent, i. e., with

its third person singular form (he), its indefiniteness (they), or, in all but (36-a), its reference to the addressee (you). That the subject of the imperative tag should be limited to you, while the other pronouns in these examples are free to agree with different features of the same antecedent, suggests that the rules governing noun-pronoun agreement must have an ad hoc restriction to limit the form of the tag subject in this context.

There is one environment in which the imperative tag subject need not be you, however (37). For one thing, there is no explicit stem subject in these sentences; equally important, there is the right combination of stem polarity and tag intonation. To permit an indefinite subject in an imperative tag, there must be either a positive stem and a 3+ tag, or a negative stem and a 31 tag. None of the other four contexts permit it.

- (37) a. Pick up that target over there, will someone?<sup>3+</sup>  
b. Put the chairs back before you leave, will everyone?<sup>3+</sup>  
c. Don't leave yet, will anyone?<sup>31</sup>  
d. \*Pick up the target over there, won't someone?<sup>31</sup>  
e. \*Don't laugh, will anyone?<sup>3+</sup>

Why indefinite tag subjects should be prohibited in contexts defined by stem polarity and tag intonation is difficult to understand, but the effect of that context is clear. And by isolating these contexts in relation to the sentences in (37), we have made it possible to write the rules necessary to generate (37-a/c) while blocking (37-d/e).

As for the impossibility of the indefinite tag subject where there is an explicit subject in the stem (38),

- (38) a. \*Someone pick up that paper over there, will someone?  
b. \*Don't anyone take any books with him, will anyone?

this can be explained in terms of the definitization rule that we saw operating in (29)-(36): pronouns having a recognizable antecedent will be definite. Since the subject of the tag is co-referential with the subject of the stem, the tag subject cannot remain indefinite. But this leaves us with the need to explain how the tag subject in (37-a/c) can remain indefinite, since one assumes that it had an antecedent in the stem subject at some point in the derivation of the sentence. It might be possible, of course, to return for an explanation of this dilemma to Sadock's WHIMPERATIVE analysis. If the tag subject in (37) was itself the subject of the stem originally, then it





We do not need the WHIMPERATIVE analysis here, and the sentences like (37) can be derived from the same source as (39).<sup>9</sup>

So far, we have found that although the derivation of the subject of the imperative tag is a fairly complicated matter, part of it depends upon the contexts defined by the polarity of the stem and the intonation of the tag. Other features that also depend on these contexts are the form of the auxiliary, the tense it can carry, and its polarity. Will, in one form or another, can occur in any tag. Can, however, in its various forms, shows up only when the tag is attached to a positive stem and has one of the two rising contours (2+ or 3+).

- (42) a. Tell me what time he's leaving, { won't you?<sup>3+</sup> (or 2+)  
can't
- b. Don't talk to me now, { will you?<sup>3+</sup> (or 2+)  
\*can
- c. Be sure to remind me about the meeting, { won't you?<sup>31</sup>  
\*can't
- d. Don't forget to call me when you get there, { will you?<sup>31</sup>  
\*can

As for the tense of the auxiliary, it must be present unless the tag has a 3+ intonation pattern, in which case, it can be past. (Notice that polarity of the stem plays no role in defining this context.)

- (43) a. Tell me what time it is, would you?<sup>3+</sup> (\*2+)
- b. Don't talk to me now, would you?<sup>3+</sup> (\*2+)
- c. \*Be sure to remind me about the meeting, wouldn't you?<sup>31</sup>
- d. \*Don't forget to call me when you get there, would you?<sup>31</sup>

And finally, the polarity of the tag must be opposite to that of the stem, except in the context in which the stem is positive and the tag intonation is a 3+ pattern. Only in this highly specific context, can the frequently discussed alternation between positive and negative polarity in the tag occur; only here can the polarity of the tag be the same as that of the stem.

- (44) a. Tell me whether he's coming, will you?<sup>3+</sup> (\*2+)
- b. \*Don't talk to me now, won't you?<sup>3+</sup> (2+)
- c. \*Tell me what time he's coming, will you?<sup>31</sup>
- d. \*Don't tell him I'm coming to his party, won't you?<sup>31</sup>

A summary of these and other facts is found in Table 1.

Another feature of imperative-plus-tag constructions that is controlled by the tag intonation and stem polarity is the possibility of the co-occurrence of the tag itself with an explicit subject in the stem. There are two contexts in which such co-occurrence does not come about: 1) when the stem is negative and the tag has a 3+ contour, and 2) when the stem is positive and the tag contour is 2+.

- (45) a. \*Don't you do that, will you?<sup>3+</sup>
- b. You keep quiet about that, will you?<sup>3+</sup>
- (46) a. Don't you do that, will you?<sup>2+</sup>
- b. \*You keep quiet about that, won't you?<sup>2+</sup>

Finally, whether the imperative-plus-tag can co-occur with please also depends on what the polarity of the stem and the intonation of the tag are. Please can follow the structure only if the intonation of the tag rises (3+ or 2+), not if it falls (31).

TABLE 1:

Summary of the effect of stem polarity and tag intonation on the imperative-plus-tag construction

INTONATION	3+		2+		31	
	POS	NEG	POS	NEG	POS	NEG
POSSIBLE AUXILIARY:						
WILL	+	+	+	+	+	+
CAN	+	-	+	-	-	-
PAST TNS POSSIBLE?	+	+	-	-	-	-
OPPOSITE POLARITY REQ?	-	+	+	+	+	+
INDEF TAG SUBJ POSSIBLE?	+	-	-	-	-	+
TAG CAN OCCUR W/OVERT STEM SUBJECT?	+	-	-	+	+	+
CAN BE FOLLOWED BY <u>PLEASE</u> ?	+	+	+	+	-	-

Proposed Derivation for the Imperative-plus-tag Construction. Now that we have a more detailed description of various imperative-plus-tag constructions that can occur, we should return once more to the question of how the structure should be derived in its various forms. One may recall that there

were two important reasons for not using Sadock's WHIMPERATIVE analysis: first, the WHIMPERATIVE is essentially a question, but we found that the imperative stem in the structure that we are studying is not; and second, the fracturing rule cannot account for structures like (47), in which there is a subject in the stem as well as in the tag.

(47) You go home now, will you?

As for the copying rule approach to the derivation of imperative tags, suggested originally by Katz and Postal, we have found that we could remedy some of what had been considered fatal defects in that type of analysis. The rules needed to account for the alternation in (48), for example, we found were also necessary for tag-questions and other constructions in which indefinite pronouns originally serving as antecedents for other pronouns are deleted.

- (48) a. Someone set up some chairs over there, won't you?
- b. Set up some chairs over there, won't someone?

We also found that these two sentences do not occur freely; the second, for example, can occur only with a positive stem and a 3+ tag, or a negative stem and a 31 tag. This restriction will have to be written into the structure description of the Imperative Deletion Rule to prevent it from applying to indefinite subjects in the context of these two stem polarity-plus-tag-intonation combinations. If the context is such that Imperative Deletion can apply, the derivation of (48-b) would follow the steps in (49); and if it is not, then the stem subject remains and the tag subject becomes definite (50).

- (49) a. Someone set up some chairs over there.      Initial Structure
- b. Someone set up some chairs over there,  
          will someone?<sup>3+</sup>                                      Subject Copying
- c. Set up some chairs over there, will  
          someone?<sup>3+</sup>    Imperative Deletion
- (50) a. Someone set up some chairs over there.      Initial Structure
- b. Someone set up some chairs over there,  
          won't someone?<sup>31</sup>                                      Subject Copying
- c. Someone set up some chairs over there,  
          won't you?<sup>31</sup>    Definitization

The first sentence in (48) cannot occur if the stem is negative with a

3+ tag, or positive with a 2+ tag. The restriction is not limited to imperatives with indefinite subjects, but prevents any co-occurrence of explicit stem subject and imperative tag (51).

- (51) a. \*Don't you talk to me, will you?<sup>3+</sup>  
b. \*No one leave, will you?<sup>3+</sup>  
c. \*You talk to someone about it, won't you?<sup>2+</sup>  
d. \*Someone tell him, won't you?<sup>2+</sup>

To prevent sentences like these, Imperative Deletion must be obligatory when either of the two combinations of stem polarity and tag intonation just described exist.

With these environmental limitations that we have been describing, it is possible to generate all the grammatical imperative tag subjects, and only those. Another facet of the tag that we can derive by copying rules is the tag polarity. If the stem polarity is [ $\alpha$ ], then in the environment in which the tag intonation is 3+ and stem polarity is positive, tag polarity will be [ $\pm \alpha$ ] to account for the alternation between positive and negative tags in that context. In all other environments, the tag polarity will be [ $- \alpha$ ].

Rules of the sort just discussed capture the essential predictability of both the subject and the polarity of the tag and treat them as formal reflexes of syntactic elements explicitly or implicitly present in the stem. But in the case of the alternating tag polarity derived by the [ $\pm \alpha$ ] rule, there is a sense of incompleteness about the analysis. The positive and negative tags here are not really in free variation with each other. Most native speakers seem to find the negative tag in sentences like (52) more polite and pleasant than the positive.

- (52) a. Stop by and see me, will you?  
b. Stop by and see me, won't you?

If a clearly syntactic element like polarity is predictable in terms of the degree of politeness intended by the speaker, the degree of politeness must be represented in the underlying structure of the sentence at the time the polarity that expresses it is generated. But if that polarity is to be generated within the context of a copying rule derivation of the imperative tag, there is a complication: the politeness expressed in the tag by polarity is expressed in the stem in other ways. Hence, what the copying rule copies

into the imperative tag in the context of a positive stem and a 3+ tag is not related to stem polarity at all. What is copied is an abstract representation of the degree of politeness, which is later transformed into tag polarity.

This complication of the concept "copying rule" is necessary with regard to tag polarity only in the limited context in which that polarity can alternate independently of the stem polarity and can, therefore, be meaningful. All other instances of tag polarity, except those in this specific environment, are completely predictable on the basis of stem polarity, generated by an [ $\alpha$ ] rule, and meaningless. But there is another syntactic reflex in the tag that is not predictable in terms of the stem--the tense. Normally the tense of the tag is present. When the tag intonation rises, however, we saw that that tense could alternate between present and past. Like the alternation in tag polarity, that of the tense seems free of constraint at first glance. But it, too, is related to the politeness of the utterance. An imperative with a past tense tag seems more polite than one in the present tense. As a result, the tense of the tag, in a context where it is not restricted to present, can be derived on the basis of the same abstract politeness marker that was used to generate tag polarity. And the need for such a marker in the derivation of both tag polarity and tag tense makes the device seem less ad hoc than it may have at first. The two rules generating the tag's polarity and tense on the basis of the politeness marker do not necessarily react to the marker in the same way, however. All four possible combinations of the polarity and tense are grammatical, and each signals a slightly different degree of politeness (53).<sup>10</sup>

- (53) a. Wait for me outside, will you?<sup>3+</sup>  
b. Wait for me outside, won't you?<sup>3+</sup>  
c. Lend me your book for a while, would you?<sup>3+</sup>  
d. Lend me your book for a while, wouldn't you?<sup>3+</sup>

When we turn to the generation of the appropriate intonation contour for the tag, we find that it, too, is based upon meaning and cannot be predicted on the basis of the form of the stem alone. Recall that when the stem is negative, a 3+ tag means to stop the action mentioned, while a 3I tag is an admonition not to start it. The stem to which these two tags can be attached (54) is potentially ambiguous between the two meanings they represent; it can be used, without a tag, in situations to which either is

appropriate. The rule generating the intonation of a tag on a particular negative stem, and thereby disambiguating the utterance, would have to be able to decipher which of the two meanings that particular stem expressed.

(54) Don't gallop here.

More difficult still, and a stumbling block to a copying approach, is the fact that the tag intonation indicates whether the speaker expects the addressee to comply with his request: a falling intonation suggests the speaker's expectation that the addressee will do what he is asked; a rising contour shows less confidence. If we attempted to generate the contour on the basis of this meaning, we would have less success than we just encountered with the Stop~Don't start dichotomy. In a stem without a tag, there is no suggestion of what the speaker expects the addressee's response to be. And if this expectation is not present in the stem, there is no way of copying it into the tag and no way of generating an intonation contour that expresses it through a copying rule approach.

The last problem, the generation of the tag auxiliary, also poses a serious obstacle to the derivation of imperative tags by copying. Neither of the auxiliaries that occur in the tag ever appears in the surface structure of the stem. This means that when either can or will is equally appropriate as the auxiliary of an imperative tag, it is not possible to predict which it will be on the basis of the stem. As Sadock (1970) remarked, any imperative stem that can take either can or will in its tag must be considered ambiguous if its underlying structure is to be the source from which those auxiliaries are copied.

Yet closer consideration will show that neither of these auxiliaries can be in the underlying structure of an imperative stem. If they were, then the imperative would be telling someone to will to do something (Bolinger, 1967) or to be able to do it. But neither of these is a legitimate command; both are essentially stative, and cannot be deliberately brought about. They can occur in the tag because there they question the will or ability of the addressee, they don't demand it. In short, the stem cannot serve as the basis for the appearance of the auxiliaries in the tag.

Our attempt to revise the copying rule approach so that it can produce the imperative tag has been somewhat successful. Both the tag subject and its polarity are basically amenable to the approach. Other elements of the

tags, based upon politeness, are less so. And still others, the expectations of the speaker regarding addressee compliance and the tag auxiliary, are incompatible with copying. Because of these last elements, we must find some other approach to the derivation of the imperative tag. The most obvious possibility seems to be an initial structure in which the tag was present. But if we do posit a deep structure tag, we must be sure that we are able to capture the generalizations concerning the tag subject, polarity, politeness, and the Stop<sup>v</sup>Don't start dichotomy that we could handle with varying degrees of success through copying rules. First, the necessary agreement between the tag subject and the second person features of the stem subject can be represented by a deep structure constraint. This same constraint can also require that the tag carry the same definiteness of its antecedent. In this way, the tag subject you in (55-a/b) would be generated at the initial structure level of the sentence by the agreement constraint, while that in (55-c) would be derived by using the definitization transformation illustrated earlier in (50).

- (55) a. You get out of here, will you?  
b. Get out of here, will you?  
c. Someone pick up that junk and heave it, will you?

As for the polarity of the tag, we have noted its total predictability, except in the context of a positive stem and a 3+ tag. We also saw that in that one context the tag polarity was linked not to the polarity of the stem, but to its degree of politeness. Somehow this relationship between this aspect of the meaning of the stem and its expression in the form of the polarity of the tag will have to be captured, and probably at the level of the initial structure. Furthermore, since the alternation between present and past tense in rising tags is also linked to politeness, it follows that whatever type of device is used to derive the tag polarity in this context can be used to generate the tense of tags also. In those contexts in which tag polarity is a reflex of that of the stem, on the other hand, this relationship is easily handled by an agreement constraint similar to the [ $\alpha$ ] rule we used earlier. If the stem polarity is alpha, that of the tag will be [- $\alpha$ ]. Or, we could use the same [ $\alpha$ ] rule as before and derive tag polarity transformationally. There seems to be little way to choose between these two approaches with regard to linking tag polarity to that of the stem.

By contrast, we noted that the stem cannot serve as the basis for predicting what auxiliary will occur in the tag. Neither will nor ability can

occur in the stem, since they cannot be the focus of a command. A particular auxiliary is simply selected on the basis of which of these factors, the addressee's will or his ability, the speaker wants to question, and this is done at the same time that other lexical items are selected, at the initial structure level.

Finally, we turn to the derivation of the tag's intonation contour, and find that it, too, must be generated in the initial structure. We have seen that this contour is itself meaningful, expressing the extent to which the speaker is confident that his request will be carried out, and/or the Stop Don't start dichotomy found in conjunction with negative stems. In addition, we have found several aspects of the tag that are dependent on the tag intonation for limitations on the form they can take. Some of these limitations affect elements of the tag that will be generated in the initial structure, e.g., polarity or tense. For both of these reasons, because it is meaningful, and because it is necessary as a context for determining the nature of certain characteristics of the tag at the initial level, tag intonation must somehow be represented there also. To do that with precision, we will have to learn exactly how the different elements of the meaning of the stem and the tag combine to produce a specific contour, and this has yet to be done. But that it must be done is clear, if we are to make use of what we have come to know regarding the effect of tag intonation and stem polarity on the form of the imperative tag.

One bonus that comes with the solution proposed here is that the presence of do in entreaties and in negative commands is explained easily and naturally. So long as tags were assumed to be generated by a copying rule, there had to be an auxiliary in the stem that was copied into the tag. This auxiliary should have precluded the presence of the substitute do from the sentence. Unfortunately for the validity of that particular approach, do occurred in the stem, while will or can appeared in the tag. How to resolve this dilemma went unanswered, as our discussion of the Katz and Postal proposal has shown; nor did Sadock's approach come any closer to a solution, since no WHIMPERATIVE, from which the imperative-plus-tag was to be derived, contained any instance of the do in question. With the derivation suggested here, however, there is no auxiliary in the imperative stem, and the same rule that provides the do for questions and statements when it is needed to carry negativity or emphasis will apply to the imperative. There is no need to



assume any longer that the do here is an oddity, "an invariable introductory formula" (Quirk, 1972), or a "fossilized main verb" (Bolinger, 1967). The fallacious assumption that there was a bond between the tag auxiliary and some comparable element in the imperative stem has been laid to rest. And the do found in imperatives is seen to be the same as the do substitute found in other sentence types.

#### CONCLUSION

To summarize, we began by noting that Katz and Postal had used the same copying rule approach to generating the imperative tag that was used in the derivation of the tag-question. But we also saw that several characteristics of the imperative tag could not be accounted for by this approach. The subject of the imperative tag was you, for example, even where the stem subject was an indefinite, like someone. Both the polarity and tense of the tag seemed able to vary without regard for that of the stem. And the auxiliary do was the only auxiliary to occur in the stem, while can and will were the only ones in the tag. In short, the form of the tag seemed to vary unpredictably and meaninglessly when compared with the stem from which its constituents were supposedly copied. Nor could Sadock's WHIMPERATIVE-based derivation bring regularity out of the seeming chaos.

In trying to find the pattern that underlay the apparent irregularity of the imperative tag, we came up with two different insights that proved helpful. First, there are contexts defined in terms of stem polarity and tag intonation to which the seemingly free variation in tag polarity and tense is restricted. Furthermore, within these contexts, the variations are not free, but linked to such semantic and pragmatic elements as the Stop^Don't start dichotomy, the degree of politeness intended by the speaker, his confidence that his request will be carried out, and whether it is the addressee's will or his ability that the speaker wants to question. The second fact that we discovered was that the apparent lack of agreement between the imperative tag subject and that of the stem could be explained by the same rules needed to derive other English structures in which a pronoun refers to an antecedent that is formally indefinite, but semantically second person. Using these insights, we attempted to revise the copying rule approach so as to make it capable of generating the various forms the imperative tag could take. Because of the important role played by semantics and pragmatics in dictating which form should occur, however, it proved necessary to assume the tag to

originate in the initial structure of the imperative-plus-tag construction, with specific constraints or transformations responsible for generating those features of the tag that are predictable on the basis of the stem. Also, since tag intonation is one of the factors defining the contexts within which the different constraints and rules apply, tag intonation itself (or the factors on which it is based) must be part of this same initial structure. What some of the factors are that dictate the form of the tag's intonation, we have already mentioned: the degree of politeness intended, the certainty of the speaker as to whether the request will be complied with, and the Stop<sup>u</sup>Don't start dichotomy. Furthermore, there are very likely others that play a role as well. For that reason, the label for the contour itself (2+, 3+, or 31) is the most precise expression, even at the initial structure level, of the tag-intonation that we have found so useful in predicting which imperative tag forms will be possible in any given sentence. All in all, what we have discovered is that the imperative tag is a highly patterned set of constructions, derived by the same rules required to derive other English structures. If it is perhaps idiosyncratic in the way in which it is controlled by the various contexts defined by stem polarity and tag intonation, the effect of those contexts on the form of the tag is systematic.

#### RELEVANCE TO ESL

How the ESL teacher will use the data provided here will be determined, of course, by many other considerations than the purely linguistic. But he can count on regularity in the polarity of the structure, with the same relationship that governed tag-questions applying here except for the one context. The teacher should also realize that the alternation between positive and negative tags in that context is meaningful. He should see the tags taking can as a subset of the whole, one whose limits are clearly defined by stem polarity and tag intonations, and the same is true of those tags permitting past tense. The different meanings associated with the various stem polarity and tag intonation combinations should help the teacher to divide the structure for purposes of presentation, and should remind him to be especially careful to avoid mixing them. It also means that students using combinations signaling a meaning other than is intended must be corrected, just as if they had selected the wrong words. But most of all, he should be alert to the many generalizations that exist to govern the various sequences, the exceptions and the need to memorize them as such, and the complexity of the overall set of tags, which suggests that this construction should be given more serious attention in ESL material.

FOOTNOTES

<sup>1</sup>IMP is an abstract marker denoting that this sentence is to be read as an imperative and not as a statement making a prediction. In this paper, the term imperative will be defined on both semantic and syntactic grounds. It expresses a command or request in a structure that has a simple uninflected verb form as its main verb. The subject may be deleted if it is the pronoun you. Questions and statements used to give commands or to make requests do not meet the formal requirements of our definition and will not be considered to be imperatives. There will, of course, be other formal grounds for identifying imperatives that will come to light as this paper progresses. Examples of imperatives are:

- a. Be quiet.
- b. You stop that.
- c. Everybody be on time tonight.
- d. Drop by sometime.
- e. Don't worry about it.
- f. Don't anyone leave.

<sup>2</sup>The conditions under which this optional polarity on the tag actually exist are somewhat limited and will be discussed later in the paper.

<sup>3</sup>Please is added to the question here to insure that it is read as a WHIMPERATIVE and not simply as an information question. The same practice is followed elsewhere in this paper.

<sup>4</sup>As for exactly what auxiliaries can occur in tags on imperatives, and what forms those can take, there is considerable disagreement from one person to another. Thorne (1966) and Arbini (1969) mention only will/won't. Bolinger (1967) also mentions can't and would. And Sadock (1970) includes these four, but adds could, couldn't, and wouldn't. Bolinger and Quirk et al. (1972) would also add the unit tag why don't you to the list of possible tags on imperatives.

<sup>5</sup>The use of the tag with negative imperatives is another controversial aspect of the whole imperative-plus-tag construction. Arbini (1969) remarks that "negative pre-verbs do not co-occur with tag-imperatives of any kind." He then suggests that sentences like Don't bring me anything, will you! or Never say that again, will you? are likely "to evoke varying degrees of dissatisfaction." Quirk et al. (1972) acknowledge that tags do sometimes occur in negative imperatives, but say that it does not happen often. Bolinger finds that "a tagged will is less likely with them (negative commands), though a sentence like Don't let them know my part in this, will you please? is possible." He explains this judgment on the grounds that he finds the "cajoling" tags and the "insistent" negative command to be essentially incompatible. Sadock also considers these sentences ungrammatical and so ignores them in developing his WHIMPERATIVE analysis. However, numerous native speakers whose thoughts are not in print seem to range from complete acceptance and an easy use of tags on negative imperatives to their total rejection. This paper will assume that since a number of natives do use tags in this context, that use must be recognized as part of English and explained.

<sup>6</sup>Thorne (1966) refers to the use of him/her to refer to the indefinite subject

in sentences like these to be "an erroneous form found among educated people in place of...your." Bolinger (1967), however, notes that "other reflexive pronouns than yourself (-selves) readily occur" in imperatives with indefinite subjects and gives several examples; and this comment generalizes easily to non-reflexive pronouns in sentences like those Thorne finds ungrammatical. Quirk (1972) and Sadock (1970) also find examples like the ones in question acceptable. In fact, Thorne is the only writer to my knowledge to take the stand he does on this point.

7 Quirk et al. (1972) flatly state that whenever there is a tag on a negative imperative stem, "the tag has a falling tone" (section 7.73). It has been my experience that a sizeable group of native speakers not only use both rising and falling intonation contours with these tags, but that they differentiate between them in the manner described here. A few have even told me that they find the rising intonation on the tag acceptable, but not the falling.

8 Neither modals nor past tense appear in the imperative stem. Sentences like (a) are not imperatives, but statements used imperatively. Unlike the true imperative, when a 3+ is attached to (a), its polarity must be the opposite of that of the stem. Further, as we noted earlier (32), the meaning is not that the addressee should cease the activity described, as it is with negative imperatives, but to avoid that activity in the future.

a. You will be on time tomorrow.

b. You will be on time tomorrow, { \*will  
won't you?

(32) a. (repeated here) You won't talk will you? ≠ Don't talk, will you?

Still another difference between the imperatively used statement and the imperative itself is that can shows up in the former regardless of the polarity of the stem or the intonation of the tag; in the latter, we found can in the tag only when the stem was positive and the tag contour rising.

And finally, although modals in the 3+ imperative tags can be past or present, this is not true of the statement used imperatively. This is another characteristic distinguishing imperatives from statements used that way.

All of this evidence is offered to show that sentences like (32) are not formally imperatives at all. Imperatives do not contain modals in their stems, only in their tags. Further, Sadock (1970) was correct in noting that no imperative without a tag conveys any sort of ambiguity as to whether can or will should be understood to have been deleted. The imperative form comparable to (a) above is simply (c), and contains no modal.

c. You be on time tomorrow.

In short, the only reason for hypothesizing that a can or a will is present in the initial structure of an imperative stem is so that it can be copied by the tag rule.

9 The derivation described to this point will not generate.

a. Get out of here, will you?

If we rely on the definitization rule to generate the [+ definite] nature of the tag subject, because its antecedent has been deleted by Imperative Deletion before, the definitization rule applies. On the other hand, we have seen that the deletion and definitization rules cannot be reversed without

making impossible a sentence like

- b. Give her a hand, will someone?

A possible solution would be to make the tag subject definite as part of the copying rule, i.e., to copy the [+ definite] feature of the antecedent. You in (a) would then be [+ definite] from the moment it was copied, while someone in (b) would be indefinite. Then if Imperative Deletion had not applied to produce (b), definitization would have been generated (c) on the basis of the antecedent, which would still be present.

- c. Someone give her a hand, will you?

<sup>10</sup> This same differentiation applies to WHIMPERATIVES, in which the tense and polarity associated with the auxiliary effect different degrees of politeness in the request.

- a. Will you wait for me outside?
- b. Won't you wait for me outside?
- c. Would you wait for me outside?
- d. Wouldn't you wait for me outside?

-----

REFERENCES

- Arbini, R. "Tag Questions and Tag Imperatives in English," *Journal of Linguistics*, 1969, pp. 205-13.
- Bolinger, D. "The Imperative in English," in *To Honor Roman Jakobson*, Vol. 1, The Hague, 1967, pp. 335-63.
- Dowling, B. T. "Vocatives and Third-person Imperatives in English," *Papers in Linguistics*, 1969, pp. 570-92.
- Katz, J., and P. Postal, *An Integrated Theory of Linguistics Descriptions*, Cambridge, Mass.: MIT Press, 1964, pp. 74-79.
- Quirk, R., et al. *A Grammar of Contemporary English*, New York: Seminar Press, 1972.
- Sadock, J. M. "WHIMPERATIVES," *Studies Presented to Robert B. Lees by His Students*, Edmonton: Linguistics Research, Inc., 1970, pp. 223-35.
- Thorne, J. P. "English Imperative Sentences," *Journal of Linguistics*, 1966, pp. 69-78.

-----

## ESL SPELLING ERRORS

Carol Chapelle  
Joan Jamieson

This study tests Simon and Simon's (1976) Phonetic Spelling Model by describing the overall patterns of spelling errors made by ESL students working on PLATO spelling and dictation lessons at the University of Illinois. The students were divided into four groups (high, low, Roman alphabet and non-Roman alphabet) and, based on Ringbom's (1977) method of classification, their spelling errors were categorized according to whether they had the same pronunciation as the target word or a different pronunciation. The errors were then analyzed to determine the proportions of error types for each of the groups of students. It was hypothesized that if the Model could be supported, there should be a difference between students' levels and the types of errors they would make, but this was not supported by the data.

### INTRODUCTION

A major problem of computer assisted instruction on PLATO, as it exists in foreign language teaching today, is its inadequate error feedback. The most obvious deficiency is at the syntactic level of analysis where a word order error is sometimes indicated when a space has not been left between two words or morphological endings have been omitted. Yet, this is certainly not the only level where errors are improperly analyzed. Morphological errors are marked as spelling errors, and conversely, spelling errors are marked as wrong words.

As easy as it is to cite the problem, its correction involves detailed and arduous analysis of actual data at all linguistic levels. Data at each level must be collected and categorized to determine patterns. These patterns would then be explained in such a way that an algorithm could be developed (Tenczar and Golden 1972) into a viable program that works better than those now in existence. Furthermore, it would be hoped that this categorization would have psychological validity in that it would yield empirical evidence to hypothesized language learning and production strategies.

Work toward such an ambitious end has been proposed by Hart (1980), and an analysis of German syntax has begun (Garrett and Hart 1981). Before these higher levels of analysis (i.e., morphological, syntactic) can hope to be perfected, however, error feedback has to be correct at the level of spelling. With this concern in mind, the purpose of the current research in English as a Second Language (ESL) is first to examine spelling errors

(in light of a phonemic-based spelling model, Simon and Simon 1976) in the hope that phonemic/graphemic correspondences will yield a pattern that can later be developed into a viable algorithm for spelling error feedback.

Toward that aim, one must ask, "When do spelling errors occur?" Reports by Brown (1970) and Simon and Simon (1976) indicate that correct spelling comes from high frequency words in the student's long-term memory. Errors occur when the word is not known. According to the Trial Spelling Method (Simon and Simon 1976), when a spelling is not known, the aural word is decoded phoneme by phoneme. These phonemes are then paired with stored phoneme/grapheme correspondences, and a trial spelling is generated. This trial spelling is matched with the student's "visual recognizer" (to determine if it looks familiar) at the level of the whole word. If so, that is how the student spells the word; if not, he tries again.

This phonemic spelling model has two underlying assumptions that are supported in the literature. First, from the basic unit of the word, phonological cues are used to generate spelling (Simon and Simon 1976, Ringbom 1977, Graham and Rudorf 1970, and Biorsky 1969). The second assumption is that there is a phoneme/grapheme correspondence in English (Hanna et al., 1966, and Cronnell 1972).

Using the word as the basic unit of analysis, one would expect to find two types of errors: one, those with the "same pronunciation," but misspelled due to multiple graphemic representations of a given phoneme, i.e., "phellow" for "fellow;" or due to the non-phonemic nature of English, i.e., "iland" for "island" (Simon and Simon 1976, Ibrahim 1978, and Ringbom 1977); and two, those with "different pronunciation" than the target word due to the speller's mishearing (Simon and Simon 1976, Graham and Rudorf 1970, and Biorsky 1969), his L<sub>1</sub> interference (Ibrahim 1978, Ringbom 1977), or due to "graphemic interference," for ESL students whose L<sub>1</sub> uses the Romanized alphabet (Oller and Ziahosseiny 1970).

The study presented here is an attempt to find empirical evidence to test this phonemic-based spelling model by discovering the nature of errors actually made by ESL students. It is also an attempt to describe overall patterns that might help point to a direction in which to begin development of a spelling algorithm based on psychological reality.



## METHOD

### Subjects

The subjects were 39 students enrolled at the Intensive English Institute at the University of Illinois, Urbana-Champaign. They were divided into two groups, High and Low level, based on their placement in the IEI. The TOEFL scores of the High group ranged from about 540-400 at the first time data was collected, to 610-460 at the second time. The range for the Low group at the first time data was collected was about 430-300, and at the second time, 520-360. They were further subdivided into two groups, Roman alphabet L<sub>1</sub> or non-Romanized L<sub>1</sub>.

### MATERIALS AND PROCEDURES

#### Defining Error Categories

Originally, Ringbom's (1977) method of classifying spelling errors was going to be applied in total because it used analysis of the whole word as its basis, classified according to same/different pronunciation (thus supporting Simon and Simon 1976), and further delineated errors according to the word's phonemes. However, several problems that were not addressed in Ringbom's article arose with this method. First of all, because the data was to be collected from the students' writing samples, whose pronunciation was to be used to determine the degree of similarity between the pronunciation of the target word and that of the form the student produced? For example, if a student typed "frut" instead of "fruit," how could it be determined whether, in the psychological system of the student, these two words had the same pronunciation? In order to apply the same evaluation to every misspelling, it was decided to focus on the target language; thus, Standard English spelling rules were used to judge erroneous forms. Therefore, the above example would be categorized as a different pronunciation error type.

Secondly, keeping a tally of specific linguistic rules for errors (Ringbom's subclassification in Ringbom 1977) was also a problem. For example, should a word with two misspelled graphemes be counted twice, i.e., "nefue" for "nephew"? Also, what should be done if one error in a word caused a different pronunciation, and another was an alternate graphemic representation, as in the case of "nefo" for "nephew"? Clearly it was necessary to modify the system to eliminate the problems identified and yet get the information needed to find the overall patterns of interest. Conse-

quently, it was decided to simplify Ringbom's system by classifying according to same or different pronunciation of the entire word (and not subclassifying).

Thus, description and analysis of overall patterns were chosen as the goal for this study so that the data could be applied to the theories and it could be determined whether looking at the word as a whole yields useful information, as Ringbom and Simon and Simon state.

### Data Collection

Errors were collected from Spelling and Dictation lessons on PLATO. In these lessons, the student hears an utterance and must type either a key word or the complete utterance, respectively. Errors were stored in a data file and were printed out when requested. Data were collected at the beginning of the semester (Time 1) and the end (Time 2).

### RESULTS

The results were obtained by comparing types of errors made by different groups of students at different times. The analysis deals with proportions of error types based on the total number of errors made for each category of interest. The results are reported in Table 1.

TABLE 1  
SUMMARY OF ERROR COMPARISONS

TIME ONE			
1.	Compare High and Low groups:	High	Low
	Proportion of different pronunciation errors	.79	.75
	Proportion of same pronunciation errors	.22	.25
2.	Compare types of errors within groups:	Same	Diff.
	*High group-Proportion of errors	.22	.78
	*Low group-Proportion of errors	.25	.75
TIME TWO			
3.	Compare High and Low groups:	High	Low
	Proportion of different pronunciation errors	.80	.77
	Proportion of same pronunciation errors	.20	.23
4.	Compare types of errors within groups:	Same	Diff.
	*High group-Proportion of errors	.20	.80
	*Low group-Proportion of errors	.23	.77

(continued)

TABLE 1 (continued)

TIMES ONE AND TWO		
5. Compare High group:	T1	T2
Proportion of same pronunciation errors	.22	.20
Proportion of different pronunciation errors	.78	.80
6. Compare Low group:	T1	T2
Proportion of same pronunciation errors	.25	.29
Proportion of different pronunciation	.75	.71
LANGUAGE GROUPS		
7. Compare errors of Spanish and non-Roman groups:	S	NR
Proportion of same pronunciation errors	.24	.21
Proportion of different pronunciation errors	.76	.79
8. Compare errors within groups	Same	Diff.
*Spanish-Proportion of errors	.24	.76
*Non-Roman-Proportion of errors	.21	.79
"GRAPHEME INTERFERENCE" ERRORS (SPANISH)		
9. Compare interference to other diff. pro. errors:	Inter	Other
*Spanish-interference and other	.08	.92
10. Compare interference errors of High and Low groups:	High	Low
Proportion of interference errors at Time 1	.16	.12
Proportion of interference errors at Time 2	.03	.04
11. Compare interference errors at Times 1 and 2	T1	T2
*High group-interference errors	.16	.03
*Low group-interference errors	.12	.04

\*Significant  $p < .01$

Note the significant ones. More errors that yield different pronunciations were found in both High and Low groups, in both Spanish and non-Roman groups at Time 1 and Time 2 (Table 1, nos. 2, 4, 8). With respect to the Spanish group, "grapheme interference" accounted for significantly fewer of the different pronunciation errors than other reasons did (no. 9). Also, at Time 2, the High and Low groups made significantly fewer grapheme interference errors than they did at Time 1 (no. 11).

#### DISCUSSION

The results of this study differed from the anticipated outcomes in several ways. First, it was hypothesized that same pronunciation errors

represent higher level problems because they are believed to result from the student's use of a more sophisticated, yet not perfect sound-grapheme correspondence system which yields predictable misspellings. Accordingly, High level students were expected to make more same pronunciation errors than the Low level students, who would be operating with a less developed system. Also, High level students were expected to make more same pronunciation errors than different pronunciation errors. To the contrary, the High and Low groups behaved similarly in every test performed on the data; both groups made more different pronunciation errors than same pronunciation errors at both times.

Also, according to Oller and Ziahasseiny's predictions, grapheme interference should have accounted for a large proportion of different pronunciation errors of Spanish speakers. Because grapheme interference errors contribute to the different pronunciation category, the Spanish group was expected to have a significantly larger proportion of different pronunciation errors than the non-Roman group. Furthermore, it was expected that Spanish students would make more different pronunciation errors that could be explained by grapheme interference than by other reasons. Neither of these predictions was supported by the data obtained (Table 1, nos. 7, 10). In fact, in Time 2 for both High and Low groups, grapheme interference accounted for almost none of the different pronunciation errors (Table 1, no. 11).

There are several possible reasons for the discrepancies cited above. First, the similarities between the data of the High and Low groups may be a result of the similarities between the two groups themselves. Both groups had pre-college entry English proficiency. Perhaps the differences between the two were not great enough to detect in this situation. On the other hand, it may have been that there were significant differences between the groups, but they were not discovered using the errors examined. For example, if the number of items tried had been collected as well as the errors made, some differences may have been discovered. Also, because the word as a unit was judged as same or different, some information was lost when there was more than one segmental error in a word.

The anticipations based on grapheme interference were too strong. Although this study found support for the existence of grapheme interference by citing cases of it, and noting a decrease from Time 1 to Time 2 (Table 1, no. 11), it was simply not a significant source of errors.

### IMPLICATIONS

From a psychological point of view, it is interesting to note that the High group made the same kinds of errors as the Low group. This would tend to support the idea that visual image rather than sounding out plays a major role in constructing the spelling, because sounding out would inevitably produce possible alternate graphemic representations. The word is either spelled correctly or incorrectly in an unpredictable way (producing different pronunciation), with no increase in sophistication of the errors. Apparently, then, as the student becomes more proficient in English, he may spell fewer words incorrectly, but he does not misspell any better.

Practically speaking, the results of this study indicate that future concentration should focus on the nature of the errors made that yield different pronunciations. It appears that alternate graphemic representations and grapheme interference do not account for enough of the errors made by anyone at any level to construct a word identification algorithm based on mappings of sound to possible graphemic representations in the native or target language.

### SUGGESTIONS FOR FUTURE RESEARCH

For the reasons suggested above, future research in this area must do two things: collect correct answers as well as errors, look at spelling errors as segments rather than (or in addition to) as whole words. Also, same pronunciation errors need not be overlooked, but patterns need to be found within the different pronunciation group, which accounts for such a large proportion of errors. Kinds of errors that might be discovered to yield different pronunciations might be mis-hearing for contrastive or other reasons, and visual confusion. When patterns are found within this category, important information will have been discovered for constructing word identification algorithms and psychological models.

-----

REFERENCES

- Boirsky, Carolyn. 1969. "Consistency of Spelling and Pronunciation Deviation of Appalachian Students." *Modern Language Journal* 53: 347-350.
- Brown, H. Douglas. 1970. "Categories of Spelling Difficulty in Speakers of English as a First and Second Language." *Journal of Verbal Learning and Verbal Behavior* 9: 232-6.
- Cronnell, Bruce. 1972. "Spelling and Sound Relations in ESL Instruction." *Language Learning* 22, No. 1: 17-27.
- Graham, R., and E. H. Rudorf. 1970. "Dialect and Spelling." *Elementary English* 67: 363-376.
- Hanna, Paul, Jean Hanna, R. Hodges, and E. H. Rudorf. 1966. "Phoneme-Grapheme Correspondences as Cues to Spelling Improvement." Washington, D. C.: U. S. Government Printing Office.
- Hart, R. 1980. "Development for Techniques for 'Intelligent' Computer Based Foreign Language Instruction." Working paper, University of Illinois.
- Hart, R., and N. Garrett. 1981. "Computer Analysis of German Interlanguage Syntax." Unpublished Ms.
- Ibrahim, Muhammed H. 1978. "Patterns in Spelling Errors." *English Language Teaching Journal* 32: 207-12. April, 1978.
- Oller, John, and S. Ziahosseiny. 1970. "The Contrastive Analysis and Spelling." *Language Learning* 21:1, 85-95.
- Ringbom, Hakan. 1977. "Spelling Errors and Foreign Language Learning Strategies." Paper presented at Conference on Contrastive Linguistics and Error Analysis.
- Simon, D., and H. Simon. 1976. "Alternative Uses of Phonemic Information in Spelling." *International Review of Educational Research* 43:1.
- Tenczar, P., and W. Golden. October, 1972. "Spelling, Word, and Concept Recognition." CERL Report x-35.

-----

EVALUATING, SELECTING AND ADAPTING PRONUNCIATION TEXTBOOKS:  
GUIDELINES FOR ESL/EFL TEACHERS

I. THE EVALUATION AND SELECTION PROCESS

Lonna J. Dickerson

This article is the first of a two-part series discussing the evaluation, selection and adaptation of textbooks which deal exclusively with pronunciation skills. The focus of Part I is the evaluation and selection process, including a thorough examination of factors leading to text selection. Part II (to appear later) deals with practical approaches and suggestions for adapting pronunciation texts to meet the needs of specific groups of language learners.

This paper is directed mainly to classroom teachers who need to provide pronunciation instruction for their students. However, much of the content is applicable beyond the domain of pronunciation teaching and would be useful to those persons who are responsible for evaluating, selecting and adapting texts for other ESL/EFL skills such as the teaching and learning of listening comprehension, grammar, reading, etc.

As we enter the 1980's, we find that two factors have created the need for a myriad of new ESL/EFL textbooks. These factors are the increasing worldwide demand for English and our changing views of what it means to teach and learn a foreign language. In trying to meet this need for varied and innovative materials, publishers are constantly bringing out new texts for all levels and all aspects of ESL/EFL instruction. Many teachers, however, are overwhelmed with the proliferation of available materials. They are finding that they do not have the expertise to do an adequate job of evaluating and selecting texts for their classes. And, after the selection process is complete, these same teachers are often faced with the problem of adapting the text to meet the needs of their students.

The tasks of evaluating, selecting and adapting texts fall upon the classroom teacher, because the teacher is in the best position to know what his/her classes need. No one else can shoulder these jobs because no one else is so strategically involved with the learners. This article is designed to help pronunciation teachers carry out these responsibilities efficiently and with a satisfying sense of having made the best decision possible. Our focus here is on the factors which enter into the evaluation

and selection process. Adaptation of texts is treated in the second part of this series.

This paper is divided into four major sections, each representing a step in the process of evaluating and selecting pronunciation textbooks. The first three steps deal with factors which should be considered *before beginning* the actual examination of texts. Steps 1-3 engage teachers in the process of thoroughly examining their unique classroom needs. After studying their teaching-learning situations, teachers should have a clear picture of the type of text which would ideally suit their students. The fourth step, then, is to locate the best match between the ideal text and available texts. This concluding step, carried out with pronunciation books in hand, deals with relevant factors to consider in the analysis of individual texts.<sup>1</sup>

#### Step 1: Sizing up the Teaching-Learning Situation

We cannot automatically label a textbook as 'good' or bad.' The issue must always be 'good or bad for *whom?*' or 'good or bad for *what purpose?*' That is, to make a wise textbook selection, our decision must take into account the textbook users and their specific situation-governed needs.

So, how can we determine the special pronunciation needs of our ESL/EFL students? This first section, Sizing Up the Teaching-Learning Situation, answers this question. Provided here is a list of eleven variables which are 'givens' and must be considered regardless of the texts or materials used in the classroom. Although these variables are not specifically textbook-related, they will guide teachers in raising questions which ultimately restrict the range of possible texts. These variables fall into three subcategories: (1) student variables, (2) teacher variables, and (3) instructional variables. The following list is not exhaustive and should be modified as necessary.

##### A. STUDENT VARIABLES

1. *Age.* Are the students adults or children? If children, are they younger children or older children?
2. *Educational Background.* Are the students literate or illiterate in their own language(s)? If literate, how many years of



formal education do they have?

3. *Language Background.* Are the students from homogeneous or heterogeneous language backgrounds?
  4. *Level of English Proficiency.*
    - a. Overall proficiency. Are the students at an elementary, intermediate or advanced level? Or are all levels mixed in the same classroom?
    - b. Pronunciation Proficiency. Do the students have serious problems (or minor problems) in understanding English and making themselves understood? Are there certain pronunciation skills which are either more difficult or less difficult? (For example, do the students have more difficulty with speaking than with comprehension, or more difficulty with stress and intonation than with consonants and vowels?)
  5. *Situations for Language Use.* Why are the students learning English? How important is pronunciation instruction for their needs? When and where will they use English? Will they use English in a non-English setting, or will they need to function in an English-speaking community? Will they use English only in informal situations, or will they use it on their jobs or in an academic setting? Will they be language users only, or will they also be language teachers? Will they use English only occasionally, or will they use it frequently? Will they use it only for a short period of time (a few weeks or months), or will they use it for a number of years or even for a lifetime?
- B. *TEACHER VARIABLES*
6. *Experience.* Are the teachers experienced or inexperienced? Have they taught pronunciation skills previously? Have they had experience in adapting textbooks or creating their own materials?
  7. *Native Language.* Are the teachers native speakers of English? If so, are their dialects very different from those represented by the textbooks under consideration? (For example, will speakers of Australian English be asked to use textbooks designed for use by speakers of American English?) If the teachers are non-native speakers, do they have pronunciation problems themselves? If so, to what extent will these problems interfere with pronunciation instruction?

8. *Preparation Time.* Will teachers have time for extensive (or minor) adaptation of materials?

C. *INSTRUCTIONAL VARIABLES*

9. *Class Time.* How much time is available for concentration on pronunciation skills? How many weeks (months) will pronunciation instruction last? How much time will be given to pronunciation each week (day)?
10. *Class Format.* Will a separate time period be devoted to pronunciation skills, or will pronunciation be integrated with other skills?
11. *Class Size.* How many students will comprise a class? By using teacher aides and/or a language laboratory, would it be possible to break the class down into smaller groups?

To help the teacher/text-evaluator focus more concretely on the above variables, it is often helpful to enter the data pertaining to specific teaching-learning situations on a checklist form. The two sample checklists shown below are illustrative of two different types of ESL programs in the United States. Group A (Figure 1) is similar to some university-level pronunciation classes, while Group B (Figure 2) is similar to many adult education programs such as those established for refugees and other new arrivals into the United States.

CHECKLIST FOR SIZING UP TEACHING-LEARNING SITUATION

Variables	Evaluation
<b>A. STUDENT VARIABLES</b>	
1. <i>Age</i> 2. <i>Educational Background</i> 3. <i>Language Background</i> 4. <i>Level of Eng. Prof.</i> <i>a. Overall proficiency</i> <i>b. Pronun. proficiency</i> 5. <i>Situations for Lg. Use</i>	young adult, university level mostly univ. grads, some high school grads heterogeneous  advanced level (can use English for communication, but many problems in all skills) comprehension: fair; speaking: difficult to understand; problems in all skill areas university studies; need to relate to native English speakers in everyday life; most will be in U.S. 2-8 yrs., but many will use English after returning to their countries; a few will be English teachers
<b>B. TEACHER VARIABLES</b>	
6. <i>Experience</i>  7. <i>Native Language</i> 8. <i>Preparation Time</i>	some ESL experience; some pronunciation teaching experience; little materials development experience  American English approximately 2 hrs. for each hour in class
<b>C. INSTRUCTIONAL VARIABLES</b>	
9. <i>Class Time</i> 10. <i>Class Format</i> 11. <i>Class Size</i>	3 hours per week for 16 weeks separate time period for pronunciation 10-15 students

Figure 1: Group A: University-Level Pronunciation Instruction

CHECKLIST FOR SIZING UP TEACHING-LEARNING SITUATION

Variables	Evaluation
<b>A. STUDENT VARIABLES</b>	
1. <i>Age</i> 2. <i>Educational Background</i> 3. <i>Language Background</i> 4. <i>Level of Eng. Profic.</i> <i>a. Overall proficiency</i> <i>b. Pronun. proficiency</i> 5. <i>Situations for Lg. Use</i>	adult varies from little to much; a few are illiterate in their native language heterogeneous beginning level beginning level living permanently in U.S.; want jobs which require English
<b>B. TEACHER VARIABLES</b>	
6. <i>Experience</i> 7. <i>Native Language</i> 8. <i>Preparation Time</i>	some ESL experience; no pronunciation-teaching experience; no materials development experience American English very little
<b>C. INSTRUCTIONAL VARIABLES</b>	
9. <i>Class Time</i> 10. <i>Class Format</i> 11. <i>Class Size</i>	five 10-minute time blocks per week in 10-week course separate time for pronunciation; can also integrate pronunciation with other skills 15-20 students; teacher aide can work with smaller groups; students may enter or exit program at any time

Figure 2: Group B: Adult Education Pronunciation Instruction

Filling out a checklist in order to assess the 'givens' of a particular teaching-learning situation is a very useful first step in the text selection process. The checklist identifies some important information about the text -- who it is for (student variables), who will teach it (teacher variables), and how it will be used (instructional variables). This information, however, says little about the content and teaching methods which are most appropriate for the learner. To address these matters, we must take the next step in text selection and evaluation: defining learning objectives. The ideal content and methodology to look for in a text will become apparent only after a list of course objectives has been developed. These objectives must clearly state the changes desired in the students' pronunciation by the end of the instructional period.

### Step 2: Defining Instructional Objectives

The second step in the evaluation and selection of pronunciation textbooks deals with this question: What should the learner be able to do that he/she cannot do now? Without clearly defined goals or objectives, our text selection process and our teaching will resemble the adventures of a man who starts a long journey into unknown territory without a roadmap or a guide of any sort. He may enjoy wandering along unfamiliar roads for a while, but it is doubtful that he will ever reach his destination. We may say that instructional objectives serve the textbook evaluator and the classroom teacher in much the same way that a roadmap serves a traveller.

However, unlike a roadmap -- which identifies the same roads, cities, etc., regardless of who uses it -- lists of instructional objectives should not be the same for all groups of students or for all teaching-learning situations. An objective which may be highly useful for one particular student may be irrelevant for another. For example, one student may need to know how to use spelling as a guide to the pronunciation of technical vocabulary, while another student may have little need to use spelling as a guide to the pronunciation of even the most common vocabulary items. Instructional objectives, then, must always be determined by keeping the student and his needs in mind.

To formulate objectives, we must keep uppermost this basic question: How do I expect my students to be different in pronunciation ability at the end of instruction from the way they are now? Said differently, what do I expect them to be able to do at the end of instruction that they cannot

do now? These questions must be answered in terms of observable behavioral changes in the students.

Figures 3 and 4 present two abbreviated sets of general objectives related to pronunciation teaching, one set for each of the two teaching-learning situations described in Figures 1 and 2 above.<sup>2</sup> Note that many of the objectives are long-term, developmental competencies which cannot be achieved easily or quickly. Also, many of the objectives reflect the integration of pronunciation with other skills such as grammar and vocabulary development. The objectives for Group A (Figure 3) are stated for a 90% level of accuracy, unless indicated otherwise.

-----  
At the end of sixteen weeks of instruction, the students should be able to:

- A. Identify major problems in their own pronunciation
  - 1. Describe problems in reception ability
  - 2. Describe problems in production ability
  - 3. Describe problems in prediction ability
- B. Use reception skills
  - 1. Discriminate English phonemes
    - a. identify segmental phonemes in phrases and short sentences
    - b. identify stress patterns
    - c. identify intonation contours
  - 2. Understand spoken English
    - a. comprehend conversation of average tempo
    - b. comprehend media other than conversation (such as lectures, radio and television, etc.)
- C. Use production skills
  - 1. Speak with pronunciation easily understood by native speakers (for b, c, d, 50% improvement in accuracy for problems identified at beginning of instruction)
    - a. speak without lengthy hesitations or rephrasing
    - b. monitor own pronunciation, making corrections when obvious mistakes occur
    - c. use correct suprasegmental patterns
    - d. use correct segmental patterns
- D. Use prediction skills where prediction is possible
  - 1. Determine correct pronunciation of segmentals
  - 2. Determine correct pronunciation of suprasegmentals

Figure 3: Partial List of Objectives for Group A  
-----

Because Groups A and B do not share the need for prediction skills and especially because Group B has so little time available for pronunci-

ation instruction, the objectives for these students are greatly reduced from those for Group A. The following list of objectives (Figure 4) is stated for a 75% overall level of accuracy, unless indicated otherwise.

-----

At the end of ten weeks of instruction, the students should be able to:

- A. Use reception skills
  - 1. Discriminate English phonemes
    - a. identify segmental phonemes as same or different
    - b. identify sentence stress
    - c. identify rising and falling intonation
  - 2. Understand spoken English
    - a. comprehend conversation about familiar topics when the native speaker is enunciating carefully
- B. Use production skills
  - 1. In the contexts listed below, speak with pronunciation which can be understood by most native English speakers
    - a. use familiar everyday phrases and sentences in conversation
    - b. read aloud the lesson material covered during instruction
  - 2. Mimic phrases and short sentences (50% improvement in accuracy for problems identified at beginning of instruction)
    - a. use correct suprasegmentals
    - b. use correct segmentals

Figure 4: Partial List of Objectives for Group B

-----

Each set of objectives outlined above will differ in two respects from those drawn up for an actual teaching-learning situation. First, the above lists are not meant to be exhaustive. That is, each actual teaching-learning situation would require additional objectives which are tailor-made for that particular group of language learners. Secondly, the above objectives are stated only at a general level. To make the most effective use of such a list, the textbook evaluator -- especially the inexperienced evaluator -- might need to subcategorize each general objective into more specific objectives, often called learning outcomes. (For example, in Figure 4, under A.1.c., it might be helpful to know that the students would be asked to discriminate rising from falling intonation in situations such as this one: Identify the intonation pattern when listening to minimally different phrase pairs.)

To determine the adequacy of a particular set of objectives, three types of questions should be raised. First, are my objectives attainable

in the actual teaching-learning situation as defined in the section above? For example, are they realistic for the language learners in question? Are they realistic in terms of the difficulty of learning new pronunciation habits and the period of time required for significant change in pronunciation ability? Are they realistic in terms of student factors such as ability level, age, motivation, language background? Are they realistic in terms of teacher factors such as native language and teaching experience? Are they realistic in terms of other instructional factors such as class size and the amount of time and emphasis given to pronunciation learning? And, if pronunciation instruction is part of a larger ESL course, are my pronunciation objectives realistic when meshed with objectives for other skills?

Secondly, is my list comprehensive? Do the students have additional needs which should be reflected in my list of objectives? On the other hand, is my list so complex or idealistic that my objectives are unattainable in the given teaching-learning situation? (Note that no objective in either Figure 3 or 4 asks for 100% accuracy.)

Finally, are the objectives defined in terms of changes in the pronunciation skills of the students -- not in terms of teacher activity or course content? The reason for focusing on change in the students is this: If we have not clearly defined the pronunciation-learning needs of our students (behavioral objectives), it is easy to focus primarily on content and teacher activity, and thus fail to select texts and teach a course which will help the students learn the specific pronunciation skills which are most critical for their needs.

To summarize, by keeping in mind the eleven variables related to the teaching-learning situation (Step 1), and by asking questions such as those in the paragraphs above, the teacher/text-evaluator should be able to define pronunciation-learning objectives which are both appropriate and attainable, thus completing Step 2 in the text selection process.

### Step 3: Determining Course Content

Step 3 asks us to use both the checklist data for the teaching-learning situation (Step 1) and the list of instructional objectives (Step 2) to determine course content. For this step, the basic questions are: What content would be most useful in helping the students achieve



the objectives outlined above? Which content areas should have major emphasis? Which areas should be omitted? Answers to questions such as these should result in a list of instructional topics. Included should be all the major divisions of pronunciation topics (stress, intonation, etc.) which are relevant for the particular teaching-learning situation and as many subtopics as the teacher/text-evaluator deems necessary or helpful for the selection process.<sup>3</sup>

Figures 5 and 6 below provide partial lists of content for Groups A and B. These lists are based upon the data summarized in Figures 1-4.

-----

A. Segmentals

1. Vowel and consonant symbols and their placement on vowel and consonant charts
2. Articulatory skills
  - a. all vowel phonemes and relevant contrasts
  - b. vowel reduction in function words and multisyllabic words
  - c. all consonant phonemes and relevant contrasts required by class
3. Prediction from spelling
  - a. vowels: all general spelling patterns
  - b. consonants: palatals (eg., *fact*, *factual*, *face*, *facial*), {Z} and {D} morphemes, contractions, y-insertion (eg., *few* /fyuw/)

B. Suprasegmentals

1. Articulatory skills
  - a. four intonation patterns: rising, falling, slightly rising, slightly falling
  - b. major stress in multisyllabic words
  - c. primary stress in constructions and utterances
  - d. sentence rhythm
2. Prediction
  - a. from spelling
    - i. all predictable stressed and unstressed vowel sounds
    - ii. all predictable consonant sounds
  - b. from spelling and syntax
    - i. the placement of predictable word stress
    - ii. the placement of utterance and construction stress
  - c. from syntax and context
    - i. the occurrence of the correct intonation contours
    - ii. the occurrence of contrastive stress

Figure 5: Partial List of Content for Group A

-----

- A. Segmentals
  - 1. Articulatory skills
    - a. all vowel phonemes and relevant contrasts
    - b. vowel reduction in function words and multisyllabic words
    - c. all consonant phonemes and relevant contrasts required by class
- B. Suprasegmentals
  - 1. Articulatory skills
    - a. two intonation patterns: rising and falling
    - b. major stress in multisyllabic words
    - c. primary stress in utterances
- C. Prediction
  - 1. The occurrence of rising and falling intonation patterns
  - 2. The placement of primary stress in utterances

Figure 6: Partial List of Content for Group B

-----

Let us note two points of comparison between the course objectives and the course content for Groups A and B. First, there is not a one-to-one match between the list of objectives (Figures 3 and 4) and the list of content for each group (Figures 5 and 6). Some objectives apply to all content areas, while others are relevant only to specific content areas. A second point to note is that many of the broad content areas, such as vowels and consonants, appear to be the same on both lists of content. However, an examination of the objectives for each group indicates that the specific goals and desired level of achievement for each group differs considerably. Also, based on the general ESL proficiency level for each group (Figures 1 and 2), we know that the articulatory instruction on vowels and consonants for Group B would have to be at a much lower level than for Group A.

We have now discussed the three steps designed to guide teachers/text evaluators through the process of clarifying their unique needs -- Sizing Up the Teaching-Learning Situation, Defining Instructional Objectives, and Determining Course Content. These steps are usually preliminary to the actual examination of pronunciation texts. When Steps 1-3 have been completed carefully, selection of an appropriate text and ultimately improved quality of learning are more like to result.

#### Step 4: Evaluating Textbooks

This section provides a two-stage procedure for the systematic examination of pronunciation texts. The first stage is intended to help

teachers/text-evaluators utilize the data from Steps 1-3 in order to narrow the range of texts to those which show the most promise of potential usefulness. The second stage deals with the assessment of potential texts in terms of a number of text-specific factors. A careful examination of these factors is crucial in order to make a wise text-selection decision.

#### A. Arriving at a Set of Potentially Useful Texts

With pronunciation textbooks in hand, Step 4 first asks teachers/text-evaluators to match the ideal characteristics of a text (reflecting the needs of individual teaching-learning situations, as determined in Steps 1-3 above) with the features of each available text. This process will eliminate from consideration those texts which are most inappropriate for the specified teaching-learning situations.

In order to identify the texts which come the closest to the ideal match, it is often helpful to use a rating system for assessing the many relevant factors. Suggested here is a system which should provide for adequate evaluation without being overly cumbersome.

- + = adequate
- ✓ = somewhat adequate
- = inadequate
- A = inadequate, but adaptation possible

To record an evaluation for each factor, teachers/text-evaluators may want to add some blank columns to the right of their teaching-learning variables, their instructional objectives and their content areas. Thus, in an easy-to-use form, an assessment can be indicated for each relevant factor for each text. An evaluation procedure such as the one below should help in the assessment process.

#### Evaluation procedure

1. Arriving at a Set of Potentially Useful Texts
  - a. Sizing up the teaching-learning situation (Step 1)
    - i. Compare each text with teaching-learning variables. Rate coverage of each variable.
    - ii. Eliminate texts which are the most inadequate. (However, at a later time, these texts may be considered as (1) texts potentially useful for adapting, or as (2) supplementary texts.)

(continued)

## Evaluation procedure (continued)

### b. Defining instructional objectives (Step 2)

- i. Compare each text with the list of instructional objectives. Rate coverage of each objective.
- ii. Eliminate texts which are the most inadequate. If no text is adequate, see if two texts could be combined. If necessary, reconsider texts eliminated in a. above.

### c. Determining course content (Step 3)

- i. Compare each text with the list of course content. Rate the coverage of each content area.
- ii. Eliminate texts which are the most inadequate. If no single text is adequate, see if two or three texts could be combined. If necessary, reconsider texts eliminated in a. and b., above.

In sum, the procedure outlined above should help teachers/text-evaluators match the requirements of different teaching-learning situations with the strengths and weaknesses of the pronunciation textbooks under consideration. By completing this procedure, they should arrive at a set of potentially useful texts.

## B. Arriving at a Final Decision

The second stage of Step 4 completes the text-selection process. In this stage, we discuss and rate those texts isolated as most promising according to the procedure in Section A above. We will consider four sets of text-related variables which do not fit into the categories listed in Steps 1-3. Unlike the variables previously discussed, these factors are not determined by the requirements of the teaching-learning situation. Rather, these are text-specific factors (such as linguistic accuracy, sequencing of exercises, format, etc.) which can be evaluated for individual texts only by an actual examination of the textbooks themselves.

In order to compare the merits of each text under consideration, we suggest using a rating system, such as the one in Section A above (+, √, -, A), combined with the following evaluation procedure.

## Evaluation procedure

### 2. Arriving at a Final Decision

- a. Evaluate each text according to the text-specific variables. Rate coverage of each variable.

(continued)

Evaluation procedure (continued)

- b. Eliminate texts which are the most inadequate. If no single text is adequate, see if two or three texts could be combined. If necessary, reconsider texts eliminated previously.
- c. Select most adequate text(s) for classroom use.

The following list of text-specific variables is not meant to be exhaustive; rather, it is meant to serve as a guide in the final stage of the text-evaluation process. Note that some of the questions listed under each variable cannot be answered apart from a consideration of the data in Steps 1-3 above. For example, to answer the first question under D. 11., 'Would a teacher's guide be helpful?', requires consideration of more than the complexities of the text itself; it also requires consideration of factors such as the teaching ability and experience of the classroom teacher.

A. *VIEW OF THE ENGLISH LANGUAGE*

1. *Linguistic Accuracy.* Are the linguistic descriptions accurate? Do they occur where needed? Are they adequate for the level of student using the text? Are they stated clearly?
2. *Sociolinguistic Authenticity.* Is pronunciation instruction viewed as an opportunity to practice appropriate and useful English phrases, sentences, dialogues, etc.? Or, does the text include archaic sentences, nonsensical sentences, portrayals of highly unlikely social situations, and sentences which, though linguistically correct, would never be used in that context by a native English speaker?
3. *Variability.* Does the text encourage the use of any variety of educated spoken English, or does it merely promote the author's own dialect, while either classifying as wrong or ignoring other equally correct pronunciations? (For examples and detailed information in support of accepting a variety of correct educated pronunciations, see W. Dickerson, 1976.)

B. *VIEW OF THE LEARNING TASK*

4. *Number of Exercises per Topic.* Is adequate practice provided for each topic covered? Are there sufficient exercises for introducing a topic during one class period and reviewing the topic with different exercises during another class hour?

5. *Exercise Length.* Are the exercises long enough? Within each exercise, is there enough material so that some items can be used for group practice and others for individual practice?
6. *Variety of Exercises.* Are there many different types of exercises? Do the exercises develop reception, production and prediction skills? Do some exercises require low-level manipulative skills while others require communicative skills? Do some production exercises require repetition or speaking while others require reading?
7. *Sequencing.*
  - a. Sequencing of exercises. Are exercises arranged along a continuum from easy to difficult? For each topic handled, are there several exercises of medium difficulty, gradually leading the learners from manipulation to communication?
  - b. Sequencing within exercises. Are items arranged along a continuum from easy to difficult?

C. *FORMAT*

8. *General Format.* Is the overall organization of the text clear? Is there an adequate table of contents and index? Overall, does the text appear interesting and attractive?
9. *Possibility for Integration with Other Materials.* Can specific lessons or exercises be used without using the entire text? Are there suggestions for points of integration with other skills, such as grammar or reading? Are there suggestions of ways to adapt the text?
10. *Homework and Other Out-of-Class Activities.* Would homework exercises be useful (particularly for the prediction skills)? If so, are they provided? Are they adequate? Are answers provided?

D. *SUPPLEMENTARY MATERIALS*

11. *Teacher's Guide.* Would a teacher's guide be helpful? If so, is it provided? Is it adequate?
12. *Diagnostic and Achievement Tests.* Would accompanying

diagnostic and achievement tests be helpful? If so, are they provided? To what extent do they reflect the objectives and content outlined in Steps 2 and 3? Are there multiple forms of each test? Are there specific instructions for administering and scoring?

13. *Tapes.* Would tape recordings be helpful? If so, are they available? Is the tape content and quality of recording satisfactory? Will students find the tapes interesting and worthwhile? Is the cost reasonable?

After completing Steps 1-4, teachers/text-evaluators should have clearly defined the pronunciation needs of their students and they should have critically evaluated how well the available pronunciation texts meet those needs. This should lead to the selection of the text or texts which are most appropriate for their specific teaching-learning situations.

#### SUMMARY

With the right text(s) in hand, classroom teachers should be better equipped to help their students achieve their ultimate goal -- a more nativelike competence in English pronunciation. To this end, we have outlined a step-by-step procedure for helping teachers define their pronunciation needs, examine available texts and then choose the materials which are most appropriate.

But, after completing Steps 1-4, what if teachers find that even the best available texts are inappropriate for their needs? Then, the only way to provide top-quality instruction is to adapt the materials they have chosen. Part II of this article will deal with suggestions and approaches for adapting pronunciation texts to meet the unique requirements of individual groups of language learners.<sup>4</sup>

-----

FOOTNOTES

- <sup>1</sup>This article assumes (1) that teachers know their teaching-learning situations well enough to apply the evaluative procedures provided, and (2) that they can identify the course objectives and content which will best meet their individual needs. Teachers whose teaching-learning situations are not clearly defined may want to use the information in this article to help them envision the scope of possible situations and thus project potential needs. Teachers who are unfamiliar with the range of possible pronunciation-teaching objectives and course content may use this article to guide their survey of potentially useful materials.
- <sup>2</sup>Objectives such as these should be useful throughout the teaching-learning process, first to help the teacher define needs in order to evaluate and select texts; secondly, to guide the instructional process by determining teaching methods, use of class time, etc.; and thirdly, to promote more effective evaluation of both the students' performance and the course itself.
- <sup>3</sup>Inexperienced teachers may need to examine a variety of pronunciation textbooks in order to determine the range of available content.
- <sup>4</sup>I wish to thank Wayne Dickerson for his many helpful criticisms and suggestions offered during the preparation of this article.

-----



REFERENCES

- Bruder, Mary Newton. 1978. Evaluation of Foreign Language Textbooks: A Simplified Procedure, Appendix II in *Adaptation in Language Teaching*. Harold S. Madsen and J. Donald Bowen. Rowley, Mass.: Newbury House Publishers, Inc.
- Chastain, Kenneth. 1976. *Developing Second-Language Skills: Theory to Practice*, 2nd ed. Chicago: Rand McNally College Publishing Co.
- Dickerson, Wayne B. 1976. Phonological Variability in Pronunciation Instruction: A Principled Approach. *TESOL Quarterly* 10:2.177-191.
- Dubin, Fraida. 1978. The Shape They Are in Now and the Shape of Textbooks to come. *On TESOL '78: EFL Policies, Programs, Practices*. H. Blatchford and Jacqueline Schachter (eds.). Washington, D.C.: TESOL, 128-133.
- Morley, Joan. 1979. Materials Development: The New Frontier, Not by Chance but by Design. *On TESOL '79, The Learner in Focus*. Carlos A. Yorio, Kyle Perkins and Jacqueline Schachter (eds.). Washington, D.C.: TESOL, 12-22.
- Tucker, C. Allen. 1975. Evaluating Beginning Textbooks. *The Art of TESOL: Selected Articles from the English Teaching Forum*, Part 2, 355-361.

-----



## A PEDAGOGICAL INTERPRETATION OF GENERATIVE PHONOLOGY

### II. THE MAIN WORD STRESS RULES OF ENGLISH

Wayne B. Dickerson

In Part I of this series, we addressed the issue of the theoretical foundations of pronunciation teaching. From this discussion, we identified the goals of pronunciation instruction and outlined the objectives by which to guide and measure the learner's progress toward the goals. Beginning with this paper, the second in the series, we turn to the practical matters of what we will teach and how we will teach it so that we can meet our goals. Because of its crucial importance in the sound shape of a word, we start with the topic of major word stress. This installment presents the complete system of word stress rules used in our ESL classes, together with the motivations for the particular form of each rule.

#### INTRODUCTION

Our consideration of the theoretical foundations of pronunciation teaching, in the first article of this series, led to this contention: ESL pronunciation teachers should strive to help their students meet two complementary goals (Dickerson 1980b). One is a *production/discrimination* goal: to be able to articulate English speech so that it is intelligible to English speakers and to be able to distinguish aurally the relevant sounds of English speech. The other is a *prediction* goal: to be able to use standard English orthography to determine which sounds and stresses should be articulated when pronouncing a word. Any learner who wishes to become a competent user of educated English must acquire both of these skills as native speakers have. The pronunciation teacher's task, therefore, is to serve as a guide so that the learner can meet these two fundamental goals most expeditiously.

One area of the sound system where the learner needs careful guidance from the teacher is word stress. The problems he has are not so much with articulation as with prediction: On which syllable does the major stress fall: *define*, *definite*, *definition*, *definitive*? The purpose of this paper is to present a pedagogical word stress system by which learners can predict the placement of major word stress.<sup>1</sup>

The topic of word stress has been chosen as the first to be discussed in this series because it is a central feature in the pronunciation of a word and therefore should be high in the priorities of instructional content. For many reasons, however, the centrality of word stress has not been appreci-

ated by pronunciation teachers or materials developers. Therefore, it might be of value to consider briefly the place of word stress in pronunciation.

The result of a misplaced stress is more than simply a mispronounced word. Like Pandora's Box, misstressing releases a variety of ills far and wide. In the microcosm of the word, the aberrant stress makes the word sound strange to the native ear and may even cause the word to be misunderstood. These problems are reason enough to treat word stress as a crucial pronunciation topic. However, the impact of misstressing reverberates far beyond the word itself. In the macrocosm of the utterance, misstressing also takes its toll. The stress of individual words is the vehicle by which the fundamental rhythm of an English utterance is conveyed. If this rhythm is not largely intact in the learner's sentences, his intelligibility will be seriously impaired. This insight constitutes a still more important motivation for teaching word stress prediction.

To stop here, however, would misrepresent the place of stress in pronunciation, because stress is not an isolated phenomenon in the sound system. There is a great deal more in the matter of stress placement than knowing where to execute a chest pulse. One major effect of stress placement is on vowel quality. For example, the presence or absence of stress determines whether a vowel will be full or reduced: *cáving*, *éxcáivate*. Furthermore, the location of stress in a word will also determine whether the stressed vowel will be long or short: *cáving*, *cávitý*. Another major effect of stress assignment is on consonant sounds. For instance, the presence of stress on a following vowel blocks the palatalization of *d*, as in *endúre*, but its absence from a following vowel promotes palatalization, as in *procéđure*. Similarly, the location of stress determines whether or not we must pronounce an unwritten /y/ in certain words. For words like *vólume*, /y/ is obligatory after /l/; for words like *volúminous*, /y/ is optional after /l/. Many other examples can be cited to illustrate the role stress plays in determining the sound segments of a word. Each example is another good reason for the learner to gain the ability to predict word stress on his own.

How have ESL teachers and textbook writers been able to help the learner handle the job of assigning stress to words? To answer this

question, we begin our discussion with a brief look at the treatment of word stress in ESL materials. What makes this topic hard to teach? This question leads us to consider the principal difficulties involved in bringing stress prediction into the ESL classroom. Have the problems been sufficiently overcome? The answer is, yes. To explain, we begin with an overview of the word stress system we have incorporated into a three-hour-per-week pronunciation course which we offer to intermediate and advanced ESL students at the university level. So, what do the rules look like? Where do they apply? How good are they? The discussion of each learner rule will answer such questions in detail.

### A SHORT HISTORY OF WORD STRESS IN ESL MATERIALS

Despite the manifest importance of word stress, this topic has never figured prominently in ESL pronunciation until very recently. The reason for this neglect has been different in different periods of ESL history.

During the taxonomic era, word stress was considered phonemic, that is, part of the given structure of a word. Therefore, like the vowels and consonants, stress had to be memorized for each vocabulary item. Pronunciation texts provided little help other than to encourage the learner to be aware of different stress patterns and to take the trouble to learn the stress of each new word. (For a survey of specific approaches to word stress in taxonomic-oriented texts, see Dickerson, 1978.) Thus, during this period, teachers and textbooks told learners essentially, 'The stress of a word is important and you must learn it. But we cannot teach it to you, because the placement of stress is unpredictable.'<sup>2</sup>

The generative era, beginning in the 1960's, quickly made this position untenable with a simple but startling claim about word stress: It is predictable. Linguists such as Chomsky, Halle, Ross, proved their point by formulating the technical rules governing the assignment of word stress. But the new insights have not changed the majority of pronunciation texts. The only thing which has changed is the reason given for not including word stress information. Now the argument is that the predictive rules are too technical to be practical, too complex to be understood, and too abstract to be relevant to speech. In effect, most teachers and textbooks are telling today's learners, 'The stress of a word is important and you must learn it. But we cannot teach it to you, be-

cause the prediction of stress is too difficult for you to learn.'<sup>3</sup>

These remarks about the present situation characterize the majority view. This view, however, is not accepted by all teachers and researchers. In fact, the sections below will show clearly that the majority position, too, is untenable.

#### PRACTICAL DIFFICULTIES WITH STRESS RULES

Judging from the complaints against generative phonology lodged by those who despair of applying its insights to teaching, we would not be far wrong to say that the most intimidating aspect of the generative system is its set of transformational rules, particularly stress rules.

Of course, the merit of these rules is that they are written in such a way that a computer could follow their explicit instructions. But the worry is that only a computer could ever use them. For any other use, such as language teaching, they appear too hard to decipher, written as they are in an algebra-like form with a distinctive-feature code which refers not to sounds but to certain characteristics of sounds. They appear too intricate to follow, with brackets embedded within brackets, requiring special conventions of use in order to maintain certain interdependencies. And they appear too complex to apply, since many words require more than one rule, each feeding the next with incomplete derivations, and each selectively drawing into consideration information about phonological structure, morphological boundaries, syntactic categories.

While the worries are real, they are not well-founded. It is true that the technical stress rules are written by specialists for specialists and not for ESL students. But the formal trappings of each rule can be unwrapped, the essential points stated in prose, and the intricacies simplified for a lay user. Thus, the assumed difficulty of rules for learners cannot stand as an adequate justification for avoiding stress rules in ESL textbooks.

The way, however, is not yet clear to introduce stress rules into ESL instruction; there are difficulties from another source. Although the recasting of rules is not easy, it is far less worrisome to the applied linguist than the *input* to the rules, the form of a word on which the rules operate. Without a proper input to the rules, no matter

how simple the rules may be, no proper output is possible.

For the technical rules, the input--the underlying phonological representation of a word--consists of a wealth of information about the nature of each vowel, the character of each consonant, the morphological constituents of the word, and its syntactic category in the string. The technical stress rules cannot operate without this information as input.

Now, consider the learner of English, for whom a simplified form of the technical rules might be devised. Could he use such rules? Yes, he could, if he knew all the information his rules needed to have about the input. May we assume that the learner comes equipped with the requisite knowledge? Definitely not. After all, he is learning the language; it is not realistic to proceed as if he has already learned it. For this reason, all research done to help the learner predict the sounds of a word must operate on this fundamental assumption: *The learner does not know the word.*

Does this basic operating assumption mean that the learner knows nothing about the input to his stress rules? By no means. He may, in fact, know some important syntactic information, such as its part of speech--noun, adjective, verb, adverb. Furthermore, with practice, he might be able to recognize certain of its morphological properties, such as prefixes and suffixes. But the fact is, he knows very little about the phonological structure of an unfamiliar word, especially about its vowels.

The learner's understandable ignorance of the vowels in novel words turns out to be such a major difficulty that it irreparably damages his chances of using current technical rules in simplified form. Dickerson (1978) stated the situation as a dilemma. Technical stress rules require that the vowel tenseness of an input word be known before the rules can assign stress. But as a practical matter, the learner does not know the word. He not only lacks information about vowel tenseness; he cannot provide such information without first knowing the stress of a word.

The dilemma is not irresolvable. There is a way to close the gap between what the learner's rules demand and what he can contribute to his rules. The gap, however, cannot be narrowed from the learner's side. The learner simply cannot supply information he does not have. Even

standard orthography by itself is almost useless in helping the learner assess vowel tenseness (Dickerson 1980a). The fact that the learner does not know the word means that he cannot accommodate the rules. If the gap is to be closed at all, the rules must accommodate the learner. Can the rules be reformulated to accept input information which the learner can realistically provide? Fortunately, the answer is, yes. Modification of the rules solves the dilemma and clears away the last major difficulty to the practical use of stress rules in the ESL classroom.

#### AN OVERVIEW OF THE PEDAGOGICAL STRESS SYSTEM

The word stress system discussed in this series consists of rules for assigning major (primary) and minor (secondary or tertiary) stresses to words. Our concern in this paper is major stress and the four main stress rules which assign major stress. Minor stress is treated at a later point in the series. We launch our study of major stress by considering in broad terms how rules assign stress to words. A detailed presentation of each rule begins with the next section.

The rules in the learner's stress system apply to the spelled form of words. To use the rules, the learner is obliged to supply from spelling the information his rules need. The most basic information required by the stress rules is the location of the *Key Syllable*. The Key is a particular syllable which stands as a reference point in a word. All four rules assign major stress in relation to the Key Syllable. The choices are limited. The rules can place the stress either on the Key or on the syllable immediately to the left of the Key, a syllable we call the *Left Syllable*.

The Key Syllable is defined in terms of spelling. A preliminary definition of the Key is *the last vowel spelling pattern of a word excluding any endings*. In the words of (1), the Key Syllables are underlined. The spelling pattern at the end of each word is written in a general way using V for vowel letter (except for a word-final *e* after a consonant letter), C for consonant letter, and # for 'end of word,' a distinction discussed below. Since spelling patterns regularly begin with a vowel letter, Keys begin with a vowel letter, too. So defined, Keys do not necessarily correspond to morphological units or to phonological or word-division syllables.



(1) <u>define</u>	<u>decree</u>	<u>interact</u>	<u>regret</u>
VCe	VV	VCC	VC#

The two middle patterns, VV and VCC, may have extra following letters, none of which affects the basic pattern. Thus, the VV pattern can be spelled with one or two extra consonant letters (*feed, peach*), and even with an additional *e* (*geese, bounce*). Each of these words is an example of the VV spelling pattern. Similarly, VCC can be spelled with extra consonant letters and an extra *e* (*tempt, badge, glimpse*). These examples illustrate the VCC spelling pattern. A more precise definition of the Key, then, includes references to extra letters: *The Key Syllable is the last spelling pattern and all extra letters at the end of a word or before an ending.*

Endings are of two types, *neutral* and *stress-governing*. Neutral endings, such as *-s, -'s, -ly, -ful, -ness*, have no effect on the placement of stress or on the assessment of vowel quality.<sup>4</sup> By contrast, stress-governing endings, such as *-es, -al, -ion*, have profound effects on both stress and vowel quality. It is therefore important to distinguish the two types of endings. The relevant criteria are simple. All neutral endings begin with a consonant letter, whereas all stress-governing endings begin with a vowel letter.

With regard to stress assignment, neutral endings are treated as invisible. Stress-governing endings, however, have such a significant impact on stress assignment that we categorize them according to their effects on stress. There are weak endings (e.g. *-es, -ed, -ing, -er, -al, -ous*), and strong sequences (e.g. *-ion, -ial, -ian, -eous*). Each of these categories will be discussed fully below in relation to the stress rules.

With regard to the assessment of vowel quality, neutral endings again play no role; they are disregarded entirely. By contrast, weak endings and strong sequences figure importantly in making predictive statements about vowel quality. For example, the presence of a weak ending must be written into vowel spelling patterns where the distinction is relevant for vowel prediction. In such a case, the symbol +W is used to mean 'before a weak ending.' For example, the VC+W pattern in the first word of (2) signals quite a different vowel quality from that signaled by the VC# pattern in the next pair of words. The +W and #

position markers are used in these patterns because they serve an important function. But the weak ending after VCC in the third column does not signal a vowel different from that found in the last two words. Therefore, +W is not written after the VCC of *planted*.

(2) <u>planes</u>	<u>plan</u>	<u>planted</u>	<u>plant</u>
	<u>plan</u> ←		<u>plant</u> '←
VC+W	VC#	VCC	VCC

When the Key Syllable has been identified, the Left Syllable can be found readily. *The Left Syllable is the vowel spelling pattern and all extra letters immediately to the left of the Key.* Defined in this way, the Left Syllable, like the Key, begins with a vowel letter. It includes all consonant letters up to the first vowel letter of the Key. In the words of (3), the Left Syllables are underlined with a wavy line. The position of a spelling pattern in the Left Syllable is important for vowel prediction in only a few cases. One such case is seen in the last word below, where we use the symbol ← to mean 'in the Left Syllable.' The leftward-pointing arrow helps us distinguish the VC in a word like *refinishes* from the VC in a word like *refines*, because the vowel quality predictions are quite different.

(3) <u>counseling</u>	<u>constrained</u>	<u>refin<u>ishes</u></u>
VV	VCC	VC←

This discussion of the Key and Left Syllable has shown that spelling patterns and types of endings are important not only for identifying the Key and Left Syllables, but also for determining vowel quality. When the learner has located the Key Syllable in a word, he has taken the first indispensable step toward predicting the sound of the word.

Information about the Key Syllable is a crucial input for the stress rules. As output, the rules will generate a stressed Key or a stressed Left Syllable. To do this, the rules operate in different ways. One kind of rule states the conditions under which stress may stay on the Key or pass to the Left Syllable. Such a rule has two potential outputs, or 'commands': Stress Key and Stress Left. Since rules of this type apply mainly to words with weak endings, we call them *Weak Stress Rules*. Another kind of rule has only a single output or command, either Stress Key or Stress Left. One of these rules assigns stress mainly to words

ending in strong sequences. For this and other reasons, to be discussed below, the rules are called *Strong* Stress Rules.

There are two Weak Stress Rules in the system. For one, the conditions governing the assignment of stress are found in the Left Syllable: Does the Left Syllable contain any part of a prefix? We call this rule the *Prefix* Weak Stress Rule. For the other Weak Stress Rule, the conditions regulating the placement of stress are contained in the Key Syllable: Is the Key spelled with a V or a VC? This is, therefore, the *V/VC* Weak Stress Rule. Since the user must evaluate the Key or Left Syllables in some way for both Weak Stress Rules, the general designation, Weak, carries the added meaning of 'evaluation required.' The terms, Prefix and V/VC, in the rule names identify the kind of evaluation needed.

Similarly, there are two Strong Stress Rules. According to one, the stress is placed, without conditions, on the Key Syllable. This is then the *Key* Strong Stress Rule. According to the other, stress must fall unequivocally on the Left Syllable. We have then the *Left* Strong Stress Rule. Again, the rule type, Strong, has added significance. Strong also means 'no evaluation required.' The terms, Key and Left, in the rules identify the syllable which receives the stress directly.

In sum, the two Weak Stress Rules are alike and the two Strong Stress Rules are alike. The two types of rules are distinguished from each other not only in their application to different types of endings and in their need for evaluation, but also in the composition of their Key Syllables. The Weak Stress Rules require one kind of Key while the Strong Stress Rules require another. The specific character of the Key is discussed with each rule.

The four word stress rules fall into the symmetrical pattern described in the matrix in (4). The abbreviations, WSR and SSR, refer respectively to the Weak Stress Rule and the Strong Stress Rule.

(4)

TYPE OF RULE

Weak: 'evaluation required'	Strong: 'no evaluation required'
-----------------------------------	--

SYLLABLE	Left	Prefix WSR	Left SSR
IN FOCUS	Key	V/VC WSR	Key SSR

This general discussion of Key and Left Syllables, neutral and stress-governing endings, Weak and Strong Stress Rules has prepared the way for a careful consideration of each rule. We look first at the Weak Stress Rules, then turn to the Strong Stress Rules.

#### WEAK STRESS RULES

The two Weak Stress Rules derive their inspiration from Chomsky and Halle's Main Stress Rule (MSR) as described in their analysis of English phonology, *The Sound Pattern of English* (1968), henceforth referred to as *SPE*. A sketch of this rule is given below in order to provide the necessary background for understanding the practical learner rules.

The Main Stress Rule is appropriately named because by its operation every word in the language receives major stress either on the Key Syllable or on the Left Syllable.<sup>5</sup> The rule 'reads' the input word and by predetermined criteria places the stress on the output. The stress placement criteria are of two kinds, phonological and morphological. They function in the rule essentially to prevent stress from passing to the Left Syllable. If the stress is blocked from the Left Syllable, it must fall on the Key.

The first blocking mechanism is the phonological shape of the Key Syllable. If the Key consists of a 'strong cluster,' the stress must stay on the Key. A strong cluster is defined as either a lax vowel followed by two or more true consonants or a tense vowel followed by any number of consonants including none at all. If the Key has any other composition, namely, a 'weak cluster'--a single lax vowel or a lax vowel followed by at most one true consonant (or a true consonant and a glide [y, w, r])--the stress will pass unobstructed to the Left Syllable. These phonological criteria for placing stress are known as the Romance Stress Rule (RSR). This rule is part of nearly every subrule in the MSR and is therefore the keystone in Chomsky and Halle's approach to stress assignment.

The second blocking device is the morphological character of the Left Syllable. The morphological constituents of a word are marked at their boundaries. The word is bounded by #, or word boundaries; constituents inside a word are marked by +, or formative boundaries. The MSR

operates within word boundaries but is not constrained by formative boundaries. A special = boundary occasionally replaces the + after a prefix formative and confines the operation of the rule to a smaller territory within the word. If the = boundary appears between the Left Syllable and the Key, stress must fall on the Key.

There is a major difference in the application and power of these two devices. The phonological device is always in effect, but it is the weaker of the two devices. In order for any word to receive stress, the phonological shape of its Key Syllable must be scrutinized. The morphological device operates only as the need arises; it is nevertheless the stronger device. Since it supersedes the phonological device, it can prevent stress from being assigned to the Left Syllable when the phonological device is unable to block such movement. Thus, while the RSR is in effect for all words, an = boundary inserted on an ad hoc basis can nullify the effect of the RSR in order to assure that stress is assigned properly.

The technical blocking devices pose two insurmountable problems for the learner, both related to the nature of the input form of a word. First, if 'the learner does not know the word,' he cannot distinguish a strong cluster from a weak cluster in the spelled form of a word. For example, in the first pair of words in (5), *eloping*, *developing*, the Key Syllables (VC+W) are identical. Yet, one Key consists of a strong cluster (a long vowel) while the other consists of a weak cluster (a lax vowel followed by one consonant). When these words are presented to the RSR, the difference in the phonological composition of their Keys results in contrasting stress patterns. The other pairs below illustrate further how unhelpful spelling is for assessing the underlying composition of the Key. Thus, the learner cannot apply the phonological criteria of the RSR for placing stress on such words. (The *ow* of *bellow* and the *y* of *rally* have underlying lax vowels which are tensed after stress has been assigned.)

- (5) *eloping*            to redefine            to bestow            to rely  
      *developing*        to determine            to bellow            to rally

The second barrier preventing the learner from using the stress rule is that he cannot know when the blocking boundary, =, has been inserted and when it has not. It is present in *to emit*, but not in *to edit*, in

to *defer* but not in *to differ*, in *excel* but not in *excellent*. Since the = boundary appears only when the RSR rule fails to block leftward movement of stress, only a knowledge of the language can isolate such failures and the need for the = boundary. But since 'the learner does not know the word,' he has no basis on which to judge the presence or absence of the special boundary. In short, the learner cannot use the morphological criterion because its presence depends on a prior knowledge of the language.

The technical rules and the nature of the input leave the learner in a difficult situation. On the one hand, the phonological criteria (tense or lax vowels in the Key) are invisible in spelling. Yet, these criteria apply consistently to all words. On the other hand, the morphological criterion (prefix in the Left Syllable) is highly visible in written words. Yet, this criterion is called upon only inconsistently. As the following discussion shows, these difficulties can be resolved. And, in a curious way, the phonological and morphological criteria turn out to be fundamental elements in the learner's rules, as well.

#### Prefix Weak Stress Rule

*The Rule.* Of the two blocking devices, Chomsky and Halle rely predominantly on the phonological one, the RSR. Their MSR gives a good idea of how successful this device is for English stress placement. The morphological device, however, appears from time to time only as a heuristic, stop-gap measure. Consequently, there is no indication in *SPE* of how reliable this device might be if used systematically to assign stress.

How useful is the morphological blocking device? From a practical point of view, the answer to this question could be of paramount importance. The learner has only one form of a word available as an input to a stress rule; that is the spelled form. Whatever criterion the stress rule applies to the input, this criterion must be clearly recognizable in spelling. The criterion of vowel tenseness has been eliminated as a possibility because such information is largely inaccessible through spelling. Prefix information, however, appears to be a better candidate because prefix shapes are well preserved in the written form of words. Could the morphological criterion serve as the primary blocking device in the learner's stress rule?

The answer is, yes. A fundamental fact about English word stress is that major stress regularly prefers stems instead of prefixes in certain classes of words. The learner's rule incorporating this basic fact is called the Prefix Weak Stress Rule (PWSR). To assign stress with this rule, the learner must respond to one question: Does the Left Syllable contain any part of a prefix? Three answers are possible: [1] There is no prefix; [2] there is a prefix; [3] there is no Left Syllable. If there is no prefix in the Left Syllable [1], stress will pass to the Left Syllable. But if there is a prefix [2] or if there is no Left Syllable [3], stress must stay on the Key Syllable. The words in (5) above are stressed by this rule and illustrate cases [1] and [2]. The learner's rule is stated as follows in (6).<sup>6</sup>

(6)

PREFIX WEAK STRESS RULE
From the Key: Stress Left, but not a Prefix. If you can't stress Left, Stress Key.

*Application.* Using the RSR, Chomsky and Halle want to show that word stress in English is controlled ultimately by phonological considerations. They avoid singling out any affix categories for special handling. They say, in fact, 'It seems, then, that there is no significant classification of affixes with respect to stress placement' (*SPE* 66). By contrast, our investigation of the morphological blocking device shows that the PWSR clearly applies to certain word domains and not to others. Suffix categories must be distinguished by the rule governing their stress.

The list in (7) contains a sample of the word domains in which the PWSR assigns major stress. Excluding the first two, each domain thereafter is identified by a weak ending as it appears on words belonging to given parts of speech. The first domain, verbs, may have no endings or any one of the endings, *-es*, *-ed*, or *-ing*. One restriction applies to this category. Verbs of three or more syllables ending in *-ate*, *-fy*, *-ize*, *-ment*, *-ute* are excluded; they are governed by other rules. The second domain includes all *ible* words. The *ible* Key may stand uninflected

or be pluralized. In the latter case, the *íbl* is the Key, and *-es* is a weak ending. In the examples of each domain, the stress is marked and the Key Syllable is underlined. The letter N follows adjectives and stands for 'noun head.'<sup>7</sup>

(7) Word Domains of the Prefix Weak Stress Rule

<u>Category</u>	<u>Part of Speech</u>	<u>Examples</u>
verbs	verbs	to <u>co</u> n <u>sí</u> der, she <u>de</u> fi <u>n</u> es, <u>y</u> éll <u>o</u> wing
ible Keys	noun, adj, adv	a <u>co</u> n <u>v</u> er <u>tí</u> ble, <u>í</u> n <u>ví</u> sible N, <u>á</u> u <u>dí</u> bly
-able	noun, adj, adv	a <u>s</u> ýll <u>á</u> ble, <u>u</u> n <u>i</u> má <u>g</u> inable N, <u>re</u> má <u>r</u> kably
-age	noun, adj	the <u>d</u> ó <u>s</u> age, <u>á</u> ve <u>r</u> age N
-al	noun	a <u>c</u> ánn <u>i</u> bal, his <u>su</u> r <u>ví</u> val
-ar	noun, adj, adv	a <u>c</u> á <u>l</u> endar, <u>p</u> ó <u>l</u> ar N, <u>c</u> í <u>r</u> c <u>u</u> larly
-ary	noun, adj, adv	a <u>l</u> á <u>p</u> idary, <u>s</u> ó <u>l</u> itary N, <u>n</u> é <u>ce</u> ssarily
-atism	noun	the <u>s</u> é <u>p</u> aratism, his <u>p</u> rá <u>g</u> matism
-atist	noun	a <u>s</u> ý <u>st</u> ematist, a <u>d</u> ó <u>g</u> matist
-ative	noun, adj, adv	a <u>p</u> re <u>s</u> er <u>v</u> ative, <u>g</u> é <u>n</u> er <u>á</u> tive N, <u>sp</u> é <u>c</u> u <u>l</u> atively
-atize	verb	to <u>a</u> ná <u>th</u> ematize, to <u>de</u> mó <u>c</u> ratize
-atory	noun, adj, adv	a <u>re</u> fó <u>r</u> matory, <u>c</u> í <u>r</u> c <u>u</u> latory N, <u>ob</u> lí <u>g</u> atorily
-ature	noun	the <u>l</u> í <u>t</u> erature, his <u>s</u> í <u>g</u> nature
-ed	adj, adv	<u>u</u> n <u>f</u> inished N, <u>re</u> p <u>e</u> atedly
-en	all words	the <u>p</u> ó <u>l</u> len, <u>u</u> n <u>e</u> ven N, to <u>m</u> ó <u>i</u> sten, <u>ó</u> p <u>e</u> nly
-er	noun, adj, adv	<u>J</u> ú <u>p</u> iter, <u>sh</u> á <u>ll</u> ower N, <u>dis</u> ó <u>r</u> derly
-ery	noun, adj	the <u>sl</u> á <u>v</u> ery, <u>sl</u> í <u>pp</u> ery N
-est	adj	<u>fl</u> á <u>tt</u> est N
-ing	noun, adj, adv	his <u>c</u> á <u>v</u> iling, <u>su</u> r <u>p</u> rí <u>s</u> ing N, <u>f</u> í <u>tt</u> ingly
-ish	2-syl adj	<u>l</u> á <u>t</u> ish N, <u>sl</u> ú <u>gg</u> ish N
-ism	2-syl noun	the <u>r</u> á <u>c</u> ism, his <u>b</u> á <u>p</u> tism
-ist	2-syl noun	a <u>l</u> é <u>ft</u> ist, the <u>fl</u> ú <u>t</u> ist
-ive	noun, adj, adv	a <u>f</u> ú <u>g</u> itive, <u>p</u> er <u>v</u> á <u>s</u> ive N, <u>ob</u> je <u>c</u> tively
-ize	2-syl verb	to <u>b</u> á <u>p</u> tize, to <u>c</u> ó <u>gn</u> ize
-or	noun, adj	the <u>b</u> á <u>ch</u> elor, <u>m</u> í <u>n</u> or N
-ory	noun, adj, adv	a <u>dir</u> é <u>c</u> tory, <u>s</u> é <u>n</u> sory N, <u>co</u> m <u>p</u> ú <u>l</u> sorily
-ure	noun	his <u>de</u> pá <u>r</u> t <u>u</u> re, the <u>f</u> ú <u>r</u> n <u>i</u> t <u>u</u> re
-y	adj, adv	<u>t</u> ú <u>s</u> socky, <u>há</u> u <u>g</u> htily

*The Key.* The learner's success using the PWSR depends in part on how accurately he identifies his starting point, the Key Syllable. To



help him, a guideline is available which specifies how Keys may be spelled. Keys in the basic--and largely Anglo-Saxon--vocabulary of English (verbs, and the *-ed*, *-en*, *-er*, *-est*, *-ing*, *-or* and *-y* groups above) are spelled with one or two vowel letters, e.g., *survivor*, *applauding*, *besieged*, *cruiser*. Keys in all other domains in (7) are also spelled with one or two vowel letters, e.g., *survival*, *laudatory*, but a qualification is in effect. Only the second vowel letter in an iV- or uV-sequence is allowed to be in the Key, e.g., *proprietary* (compare *besieged*), *intuitive* (compare *cruiser*). The guideline is given in (8).

(8) The Key in Weak Stress Rule Words

The Key is spelled with one or two vowel letters. However, only the V of iV- and uV-sequences is permitted in the Key.

In addition to the main guideline in (8), the learner finds in each lesson pointers about the Key in words with new endings. This is important information when endings change form because of added endings. For example, *-y* changes to *-i-* in *-ier*, *-iest*, *-ily* sequences and in *-aries*, *-arily* sequences.

*Predictive Accuracy.* The twenty-eight categories above represent no fewer than 22,000 words in contemporary English.<sup>8</sup> Depending on the domain, the PWSR will generate stress predictions with an accuracy of 95% to 99%. This level of correct prediction equals that of the MSR for the same word classes, if we disallow the use of the ad hoc boundary for exceptions. Of course, the sets of exceptional-stress words are different for the two rules because different criteria are used to assign stress.

Since word stress can be predicted so successfully with the PWSR, it is clearly worthwhile for the student not only to learn to use the rule but also to memorize the few exceptions.

*Preparation.* To use the PWSR, the learner does not have to know the word he is stressing. But he must be able to supply some crucial information about the word. [1] The part of speech and [2] the kind of suffix, if any, are necessary clues for determining which rule applies to the word. If the PWSR applies, [3] the location of the Key Syllable and [4] the presence or absence of a prefix are important in order to operate the rule.

For intermediate-level learners, the ability to judge part of speech

--noun, adjective, verb, adverb--is a skill they bring to the pronunciation class. Formal instruction in their grammar classes can be depended on to provide the requisite training. The ability to recognize affixes and to identify the Key, however, are skills the learner must gain in the pronunciation class, because nowhere else in the typical ESL curriculum are these topics dealt with as thoroughly as necessary. Preparation in these areas is not difficult, because the learning tasks are divided into carefully graded and highly redundant segments and spread over several weeks of instruction. The learner is gradually introduced to a set of forty Anglo-Saxon and Latinate prefixes (the only ones to affect stress assignment). His task is to learn to recognize prefix shapes in spelling; he is not responsible for prefix meanings. He becomes familiar with the location of the Key as he studies the vowel spelling patterns he will need in order to predict the vowel sounds of stressed and unstressed syllables. Endings, however, are introduced in small groups as the stress lessons are taken up. In our experience, the preparatory work on prefixes and vowel spellings (and Key Syllable) usually takes about six weeks (out of a sixteen-week semester) during which time articulatory work of all kinds proceeds. Thereafter, the stress rules are introduced along with their word domains.

It is remarkable how quickly students can use the rule to assign stress. And with their facility comes a streamlining; they rarely remember the wording of the PWSR past the first week. The rule is reduced to its central question: Is any part of a prefix in the Left Syllable? Even the name of the rule becomes simply the Prefix Rule.

#### V/VC Weak Stress Rule

*The Rule.* There are certain categories of words in which few prefixes appear or in which prefixes pose no obstacle to the leftward movement of stress. In these particular categories, the PWSR is of little value. The morphological stress assignment criterion clearly must be replaced by one that is more relevant. What factor, visible in spelling, might have predictive value?

A crucial observation about these words is that the spelling of their Key Syllables is a surprisingly reliable guide to their phonological structure, unlike the spellings in the Keys of PWSR words. With this observation, we return to the possibility that the phonological criteria

of the RSR might be useful after all. To examine this possibility, two questions must be answered: How do spellings match with weak and strong clusters? How does stress behave with respect to these spellings?

Spelling patterns correspond to weak and strong clusters as shown in (9). The VC(+W), VCC and VV patterns have been introduced above. The V pattern, however, is new to the discussion and merits some comment. This spelling consists of a single vowel letter with no following consonant; in *manual* and *continuous*, the V stands alone before the *-al* and *-ous* endings.<sup>9</sup> With respect to stress, these spellings act like weak and strong clusters. Word stress is blocked by a Key spelled VCC or VV, but passes unhindered to the left of a Key spelled V or VC. Thus, in select word domains, we see what appears to be the operation of the RSR.

(9)	Weak Clusters		Strong Clusters
	V                  VC		VCC                VV

Given the above findings, there seems to be ample motivation to formulate a learner rule which mirrors the technical rule. The pedagogical counterpart of the RSR is called the V/VC (vee-vee-cee) Weak Stress Rule (VWSR). It asks for an evaluation of the Key Syllable: Does the Key contain a V or a VC? If so, the stress falls on the Left Syllable. If there is no Left Syllable, or if there is no V or VC in the Key, stress must stay on the Key. The VWSR is given in (10).<sup>10</sup>

(10)	V/VC WEAK STRESS RULE
	From a V or VC Key: Stress Left. If you can't stress Left, Stress Key.

*Application.* The VWSR is limited in scope by comparison with the PWSR. As the list in (11) reveals, the V/VC Rule applies principally to words with weak endings. In one case, however, there is no ending: words having a final *ic*. In such words, *ic* is the Key Syllable, as in *acoustic<sub>ic</sub>* (like *acoustical*). One category below is especially interesting. As discussed in Dickerson 1977a, it seems that words ending in *-ant*, *-ance*, *-ancy*, *-ent*, *-ence*, *-ency* have been undergoing a transition from a pre-

dominantly morphologically-controlled pattern (the PWSR) to a predominantly phonologically-controlled pattern (the VWSR). At this moment, the vast majority of words are governed by the VWSR. There are, however, many words which may be stressed either on the Key or on the Left Syllable, suggesting that the change is still underway.

The sample categories in (11) below are defined by ending and by part of speech. The example words are stressed and their Keys underlined. As before, N refers to 'noun head,' and marks a preceding adjective.<sup>11</sup>

(11) Word Domains of the V/VC Weak Stress Rule

<u>Category</u>	<u>Part of Speech</u>	<u>Examples</u>
ic Keys	all words	a mechán <u>ic</u> , clássi <u>c</u> N, to pícn <u>ic</u>
-al	verb, adj, adv	to tó <u>tal</u> , monum <u>ental</u> N, analýti <u>cally</u>
-an	noun, adj, adv	a tobó <u>ggan</u> , Améri <u>c</u> an N, hú <u>manly</u>
-ance	noun	his tó <u>lerance</u> , an alló <u>wance</u>
-ancy	noun	the irr <u>élevancy</u> , his flambó <u>yancy</u>
-ant	noun, adj, adv	an accó <u>untant</u> , abú <u>ndant</u> N, míl <u>itantly</u>
-ence	noun	a r <u>ésidence</u> , the conv <u>érgence</u>
-ency	noun	his consti <u>t</u> uency, an em <u>érgency</u>
-ent	noun, adj, adv	the pr <u>ésident</u> , díff <u>erent</u> N, consi <u>stently</u>
-is	noun	the g <u>énesis</u> , a synó <u>psis</u>
-oid	noun	a trá <u>pezoid</u> , an ellí <u>psoid</u>
-on	noun, adj, adv	the phlogi <u>ston</u> , crí <u>ms</u> on N, c <u>om</u> monly
-ous	adj, adv	g <u>énerous</u> N, trem <u>é</u> ndously
-um	noun	a conti <u>nuum</u> , the refer <u>é</u> ndum
-us	noun	an é <u>xod</u> us, the thes <u>á</u> urus

*The Key.* The Key in VWSR words is spelled like the Key in PWSR words which are not of Anglo-Saxon origin. Therefore, the full guideline given in (8) above characterizes the Key in all word groups of (11). Thus, the learner has a principled basis on which to distinguish the VV Keys of *jealous* and *complainant* from the VC+W Keys of *fortuious* and *zodiacal*.

*Predictive Accuracy.* The word sets in the above list represent approximately 11,000 words in modern English. The VWSR will provide proper stress for about 95% of these. The MSR will do somewhat better. The reason for the difference is that the technical rule operates with vowel qualities given in the input words, while the pedagogical rule depends

on the spelled reflection of these vowel qualities, a reflection which is not always perfect.

Both rules have trouble with the *-Vnt*, *-Vnce*, and *-Vncy* word sets because of the 'transition factor' discussed above. Our rule generates correct predictions 96% of the time. Both rules are highly successful with the *ic*, *-al*, *-an*, *-on*, and *-ous* categories, where the VWSR assigns stress accurately about 98% of the time. The MSR has the edge in the *-is* set because its input has tense vowels in three large subsets which are exceptions for the VWSR: *-ésis* (but not *gênesis*), *-ítis*, *-ósis*. With a knowledge of these three sequences together with the VWSR, the learner can predict proper stress for 98% of the *-is* words. As for the *-oid*, *-um* and *-us* sets, our rule assigns stress with an accuracy of about 95%.

Despite a predictive failure of about 5% overall, the VWSR is a highly useful rule for the learner. Fortunately, most of the exceptions fall into neat groups which make their learning easier.

*Preparation.* The first step toward using the VWSR is to recognize which word domains are controlled by the rule. To take this step, the learner must be able to [1] identify parts of speech and [2] recognize endings. As the list in (11) shows, the combination of these two types of information defines where the VWSR applies. The second step is to decide where to place the stress according to the VWSR. The learner can take this step only if he is able to [3] identify the Key Syllable and [4] determine the spelling pattern of the Key.

The skills which the learner must have in order to use the VWSR were discussed above in connection with the PWSR. As noted, the learner becomes familiar with the Key Syllable in his study of vowel spellings. Now, his skill in recognizing the V, VC(+W), VCC and VV patterns not only acquaints him with the Key, but also contributes directly to his ability to use the stress rule. Thus, the learner's preparation for the PWSR is also sufficient preparation for the VWSR.

The V/VC Rule itself is not difficult for the learner to apply. His abilities to identify the Key and to judge spellings come together readily. In fact, the learner becomes so familiar with the spelling patterns involved that he quickly aligns the patterns with stress commands:

V and VC mean 'Stress Left,' VCC and VV mean 'Stress Key.' By comparison, the learner finds more of a challenge in remembering when to apply the rule. Adequate practice, however, helps him meet this challenge.

The Weak Stress Rules represent solutions to two prediction problems arising out of this basic fact: The learner does not know the word. First, the technical MSR assigns stress on the basis of underlying vowel qualities. No matter how simply the technical rule might be restated, the learner does not know the necessary vowel qualities on which the rule can operate to stress a novel word. The Weak Stress Rules solve this problem; they permit the learner to predict the stress of words without knowing anything about vowel quality. Second, to pronounce a word, the learner must know something about the quality of its vowels. The technical system cannot help him, because, in it, vowel qualities are given, not predicted. Nor can he turn to standard orthography which, without stress information, is hopelessly inadequate to signal vowel qualities. The Weak Stress Rules solve this problem, as well. As discussed later in this series, the learner can predict most of the vowels in a word by adding his newly-generated stress information to available spelling information. Thus, when the learner uses the Weak Stress Rules, he gains a major handle on the prediction of how a word should sound.

#### STRONG STRESS RULES

The Strong Stress Rules assign stress without regard to the phonological or morphological characteristics of a word. For this reason, the learner need not consider these factors when applying the rules. In keeping with our working assumption that the learner does not know the word he is stressing, we have designed the strong rules so that the learner can proceed mechanically to the proper stress. To show how these rules work, we begin our discussion with the Key Strong Stress Rule and then move to the Left Strong Stress Rule.

##### Key Strong Stress Rule

*The Rule.* Long before generative phonology appeared on the scene, the observation had been made that words ending in *-ion*, *-ial*, *-ious*; and *-ian* are regularly stressed on the syllable immediately before the ending: *distráction*, *colónial*, *impérvious*, *comédian*. Chomsky and Halle's contribution was to show that these endings are not unique in any way but fit

comfortably into the general pattern of word stress captured by the MSR. Consequently, there is no special category in *SPE* for these endings. Instead, they are accommodated here and there throughout the MSR according to their phonological and morphological properties.

For pedagogical purposes, however, there are compelling reasons to treat these endings as a special group. And, of course, if the group exists, a new rule must also exist to state the special stress characteristics of the group. Why segregate these endings and involve a new rule? In the first place, the endings form a highly visible set. Not only do many members of the set occur with high frequency, but each member conforms to a single, easily recognized shape, an *i* followed directly by another vowel letter. These factors make it possible for the learner to identify the endings by type without having to memorize each individual token. In the second place, nearly all words having *iV*-endings behave alike with respect to stress (and vowel quality). This powerful and useful observation would be lost if the endings were incorporated into the general stress patterns of the Weak Stress Rules. Furthermore, to stress these words by the Weak Stress Rules would require needless syllable evaluation, a complication when compared with the straightforward assignment of stress to the syllable left of the ending. In the third place, regardless of what is done for stress purposes, these endings must be isolated because they are involved in special vowel quality patterns and consonant changes. (These regularities will be considered in later installments of this series.) If the endings are identified for stress purposes, it is easier to make simple statements about the vowel and consonant patterns, as well. Thus, the overall pedagogical efficacy of segregating the set of *iV*-endings from all others is the primary motivation for developing a rule to assign stress to the set.

The present shape of the stress rule is the result of a series of extensions and refinements, each one of which has increased the generality of the rule. First, in keeping with our definition of the Key Syllable as the vowel spelling pattern and all extra letters preceding an ending, the stressed syllable in words like *distráction*, *colónial*, *impérvious*, *comedian* is called the Key. By labeling the position left of the ending as the Key, we can state the stress and vowel quality rules in an economical way. Second, what is true of *-ion*, *-ial*, *-ious* and *-ian* also holds

for many other, less frequent endings. Thus, by generalizing the category of endings to iV-endings, a broad range of vocabulary is shown to be subject to the same general patterns. Third, the special category of iV-endings does not include the strings *-ies*, *-ied*, *-ier*, *-iest*, or *iet#*, as evidenced by different behaviors with respect to stress and vowel quality. To draw this distinction more clearly, the category of iV-endings was renamed 'strong' iV-endings. The term, strong, reflects the fact that stress invariably falls on the Key regardless of its composition or the composition of the Left Syllable. Fourth, the patterns of stress and vowel quality occur not only at the end of words but also deeply inside words. The designation 'ending' is therefore not appropriate for iV-strings word medially, as in the words *stationary*, *behaviorism*. For this reason, the term 'sequence' is preferred over the term 'ending,' because it is neutral as to position. Thus, the rubric, 'strong iV-sequences,' helps us to see more clearly the application of the rule in a wider range of environments. Fifth and finally, continuing research has shown that the patterns of stress and vowel quality, so obvious in words with strong iV-sequences, also characterize words with eV- and yV-sequences, as in *extraneous*, *ichthyosaur*. This observation led to the final refinement and extension of the category; it is now designated simply as 'strong sequences' to incorporate all three types of sequences.

The Key Strong Stress Rule (KSSR), as stated in (12), reflects the evolutionary process outlined above.

(12)

KEY STRONG STRESS RULE
For Strong Sequences: Stress the Key Syllable.

*Application.* The domains of the KSSR are central to the rationale and development of the rule and were therefore discussed above. Here, we illustrate the strong iV-, eV- and yV-sequences. It should be noted that a strong sequence alone defines the domain of the stress rule; part of speech is irrelevant if a word contains a strong sequence.

Among the three types of strong sequences, the iV type is the most frequent. It comes in a myriad varieties, only some of which are illustrated below. By comparison, the eV-sequences are highly restricted.



First, there are only ten word-final eV-sequences: *-ea*, *-eal*, *-ean*, *-ear*, *-eate*, *-eo*, *-eon*, *-eous*, *-eum*, *-eus*. Second, these strings are strong sequences only if no prefix appears to the left of them, e.g. *núclear* vs. *uncléar*. When word medial, however, eV-sequences are unrestricted. The yV-sequences occur only word medially and are exemplified below left of the major stress. As shown later in this series, the KSSR applies to strong sequences in this position but for purposes of assigning minor stress.

The three domains of the KSSR are illustrated in (13).<sup>12</sup> In each case, the Key Syllable to the left of the strong sequence is underlined.

(13) Word Domains of the Key Strong Stress Rule

<u>Category</u>	<u>Examples</u>
Strong iV-Sequences	famíliar, Arábian, demóniac, rádiate, média, trívial, fóliage, William, áviary, brilliant, déviance, ámiable, mýriad, Virgínia, láriat, álien, obédience, grádiént, lénieney, ídiom, cháriot, audítion, sénior, stúdio, grácious, pódium, Július
Strong eV-Sequences	área, córneal, ócean, línear, clýpeate, ródeo, pígeon, petróleum, cadúceus, hídeous
Strong yV-Sequences	èmbryólogist, ìchthyólogical

*The Key.* The potential shapes of the Key in strong-rule words contrast markedly with those in weak-rule words. For this reason, the learner needs a guideline to help him distinguish Keys which are permissible from those that are not. The guideline given in (14) will show him clearly that *au* is the Key in *nduseous*, but that *ua* is not the Key in *situátiön*. Similarly, he can tell that *Aléutian* has a VV Key but that *influéntial* does not. He will also know that *coáxial* does not have the VV spelling of *coax*.

(14) The Key in Strong Stress Rule Words

The Key is spelled with only one vowel letter.  
However, *au*, *eu* and *ou* are permitted in the Key.

*Predictive Accuracy.* Despite the small number of domains in which the KSSR operates, the number of words represented is on the order of 10,000. This includes words with final as well as medial strong sequences, as

illustrated above. A large portion of this vocabulary is common in educated speech; the learner cannot avoid extensive contact with it. To begin to meet this challenge, he needs a good stress rule.

The KSSR is the right rule; it is simple and accurate. Even though it is the simplest of the four main stress rules in our pedagogical system, the KSSR places stress correctly on more than 99% of the words it applies to. Even the few exceptions which exist, such as *denial*, *appliance*, *classifiable*, *idée*, *muséum*, *Européan*, fall into easily-recognized sub-patterns.

*Preparation.* The KSSR makes only minimal demands on the learner. He must be able to [1] recognize strong sequences and [2] identify the Key Syllable. The first skill is easily acquired. By excluding *ie#*, *ies*, *ied*, *ier*, and *iest* strings, the learner can be sure that all other iV-sequences are strong. To these sequences, he must add all yV-sequences and a limited number of eV-sequences. The notion of the Key is familiar to the learner from other stress rules and his work with vowel spellings. How well he can locate the Key depends on how accurately he can isolate strong sequences. Our experience shows that the learner's ability to recognize strong sequences and identify the Key Syllable grows rapidly with a modest amount of practice.

#### Left Strong Stress Rule

*The Rule.* Unstressed by the foregoing rules is a large corpus of words which fall into four categories. To appreciate the pedagogical rule which accommodates these words, it is helpful first to consider the characteristics of each category and the way each is treated in the technical analysis.

In the first category, there are words of three or more syllables which do not end in a weak ending, a strong sequence or any highly recurrent syllable. In addition, these words have lax vowels in the ultimate syllable. The words, *máverick*, *apócalypse*, *strátegem*, *ténebrith*, are examples. In words like these, Chomsky and Halle omit the final syllable and apply their RSR to the penultimate syllable. In nearly all such words, there is a weak cluster in the penult, so that the RSR assigns stress to the antepenultimate syllable. In our terms, the penultimate syllable is the Key and the stressed, antepenultimate syllable is the Left Syllable. Supporting this analysis of Key and Left Syllables is the fact

that the vowels in the Left Syllable have the qualities expected of vowels in Left Syllables, as established by independent criteria.<sup>13</sup>

The second category of words in which stress remains unassigned so far consists of nouns of three or more syllables and ending in *y*, such as *industry*, *company*.<sup>14</sup> These words cannot be stressed satisfactorily by any one of the rules previously discussed. Chomsky and Halle treat the *y* as a glide and not as a vowel. Thus, in the words *industry* and *company*, for example, the syllable first evaluated by the RSR is respectively, *-ustry* and *-any*, each of which has only one vowel. Two steps are required to place stress properly on the word *industry*. First, one subpart of the MSR assigns stress to the *-ustry* because it contains a strong cluster (a lax vowel followed by two true consonants). Then another subpart of the MSR (the Stressed Syllable Rule) moves the stress one syllable to the left: *îndustry*. For the word *company*, only one step is necessary for stress assignment. The MSR finds a weak cluster in *-any* (a lax vowel followed by only one true consonant and a glide), and therefore allows stress to pass to the left: *çompany*. The crucial observation about the rather involved procedure just described is that the stress of nouns ending in *y* does not stay on the syllable first evaluated (the Key) but comes to rest to the left of that syllable (on the Left Syllable). Again, the designation of Key and Left Syllables in these words is confirmed by an analysis of vowel qualities.

The third category of words yet to be stressed also contains words of three or more syllables. These words end either in a tense vowel (unlike the first group of words above) or in a monosyllabic formative of Greek origin. In this group are words like *ênvelope*, *çompensate*, *démocrat*, *ênnergize* (*ênnergise*). In the technical analysis, these words first receive stress by the MSR on the final syllable, either because that syllable contains a strong cluster (a tense vowel) or because the monosyllabic formative is preceded by a blocking boundary. Then, all of these words pass from the MSR to the Alternating Stress Rule (ASR). The ASR moves the major stress to the antepenultimate syllable, but not by syllable evaluation. The rule simply 'counts' two syllables to the left of the final syllable and reassigns the stress. Significantly, the newly-stressed syllable quite uniformly has the vowel quality expected of a Left Syllable vowel. Thus, if the antepenultimate syllable is the Left Syllable, the penultimate syllable is the Key.

The fourth category of words so far unstressed is very large. It consists of two-syllable nouns such as *a c amel*, *a c ity*, *a pr oduct*, *a c onvoy*, none of which have any special endings. The bulk of two-syllable nouns are stressed on the penultimate syllable. The generative analysis yields this result after either a one-stage or a two-stage (assignment and reassignment) process. Once more, vowel qualities in the stressed syllable suggest that the stress is on the Left Syllable. Thus, the final syllable is the Key.

The four categories just described share this basic feature: All their words are stressed on the Left Syllable. As noted, there are many different ways stress may come to the Left Syllable. But because of the technical complexities, no pedagogical rule could successfully replicate these various routes. Fortunately, such a replication is not required. Complexities can be disregarded and stress assigned directly to the Left Syllable by a single, simple learner rule. To understand the rule, however, it is necessary now to define some useful terms.

Crucial for stress assignment is the identification of the Key Syllable. To begin with, how can the learner find the Key in the diverse set of words under discussion? Consider again words from the first three categories above, such as *m averick*, *industry*, * envelope*, *d emocrat*. In addition to having stress on the Left Syllable, these items are alike in two other important respects. First, they consist of three or more syllables in their uninflected form. For this reason, we call them *long* words. Second, the Key is not the last spelling pattern in these words; the Key is in the next-to-the-last spelling pattern. To find the Key in these long words, the learner must treat the last spelling pattern like he treats weak endings and strong sequences, namely, as a guide to the Key. Since the last spelling pattern is not an ending or a sequence, we refer to it as a *terminal*. The designation, terminal, is reserved for the last spelling pattern of long words which are not accommodated by the other three stress rules. So, where is the Key in the words? It is the spelling pattern and all extra letters immediately to the left of the terminal.

Now, consider the fourth category discussed above, namely, two-syllable nouns like *c amel*, *pr oduct*. By slightly extending our terminology for word length, we can refer to this category as *short* nouns. 'Short'

identifies a word which has only two syllables in its uninflected form. For short nouns, the Key is the last spelling pattern and all extra letters at the end of the word.

The newly defined terms, long, short, and terminal, allow us to state the Left Strong Stress Rule (LSSR) in a way which is simple yet powerful. The rule, given in (15), places stress on the Left Syllable of words with terminals and on the Left Syllable of short nouns.

(15)

LEFT STRONG STRESS RULE
For Terminals and Short Nouns: Stress the Left Syllable.

As framed, the LSSR is much more powerful than its technical counterpart, the ASR. It is more powerful because it incorporates not only the words which are stressed by the ASR according to *SPE*, but also a large number of other words from the MSR. With respect to stress, these latter words behave precisely the same as the ASR words.<sup>15</sup>

*Application.* The LSSR applies to two large domains, words with terminals, and short nouns. Short nouns are illustrated at the end of the list in (16); they will be discussed more fully below. Long words with terminals, however, require comment here. First, some terminals recur frequently; these are given in the first nine lines of the list. Less frequent terminals are illustrated under 'Other Terminals' according to part of speech. Second, certain conditions refine the domains and even affect the rule which applies. Only three refinements are mentioned here. The first concerns all words with terminals: If a strong sequence occurs immediately to the left of the terminal, the KSSR applies. Thus, in words like *bactériostat*, *chôreograph*, *ichthyosaur*, *unionist*, the Key is left of the strong sequence. The second refinement concerns long verbs: They must have no prefix immediately to the left of the terminal. This refinement allows the learner to distinguish words like *to reinstâte* (PWSR) from words like *to reinstigate* (LSSR).<sup>16</sup> The third refinement concerns only words ending in *-ize*, *-ist*, *-ism*, *-ish*. Such words require the LSSR only if they have no weak ending (or strong sequence) immediately to the left of the terminal. Words like *fêverish*, *pôsitivism*, *câpitalize*, which contain internal weak endings, are stressed according to the rule demanded by the weak ending in question.

Although the list below segments domains into numerous subcategories, there are only two relevant domains, a long word with a terminal, and a short noun. The examples in the list have underlined Keys, and the letter N, for 'noun head,' follows the adjectives.<sup>17</sup>

(16) Word Domains of the Left Strong Stress Rule

<u>Category</u>	<u>Part of Speech</u>	<u>Examples</u>
-acy (a unit)	long noun	your áccur <u>acy</u> , his cándid <u>acy</u>
-ate	long words	to áll <u>ocate</u> , some ch <u>ocolate</u> , íntim <u>ate</u> N
-ish	long adj	dév <u>ilish</u> N, yéll <u>owish</u> N
-ism	long noun	the det <u>erminism</u> , her óptim <u>ism</u>
-ist	long noun	a mon <u>opolist</u> , a psych <u>ologist</u>
-ize (-ise)	long verb	to w <u>esternize</u> , to m <u>erchandise</u>
-ute	long words	to c <u>onstitute</u> , a p <u>arachute</u> , d <u>estitute</u> N
-y	long noun	the ge <u>ography</u> , the univ <u>ersity</u>
-y	long <i>fy</i> words	to m <u>agnify</u> , an ámpl <u>ifier</u> , s <u>atisfying</u> N
Other Terminals	long noun	an ácro <u>bat</u> , a b <u>oomerang</u> , a c <u>entipede</u> , a ph <u>onograph</u> , a t <u>elegram</u> , a únic <u>orn</u>
Other Terminals	long verb	to áuto <u>graph</u> , to flább <u>ergast</u> , to p <u>antomime</u> , to r <u>idicule</u> , to s <u>acrifice</u> , to v <u>ivisect</u>
Other Terminals	long adj	bárit <u>one</u> N, dérel <u>ict</u> N, génu <u>ine</u> N, óppos <u>ite</u> N, páral <u>lel</u> N, t <u>aciturn</u> N
Short Noun	short noun	a búzz <u>ard</u> , a fáuc <u>et</u> , a lán <u>tern</u> , a r <u>ecord</u> , a ré <u>bel</u> , a p <u>ygmy</u>

*The Key.* As described above in connection with the KSSR, the learner's guideline identifying the Key in strong rule words permits only a single vowel letter in the Key except that *au*, *eu*, and *ou* form possible Keys. This guideline holds for the LSSR, as well, but applies only to words with terminals, not to short nouns. For words with terminals, the learner knows where the Key is when an *iV-*, *eV-* or *yV-* sequence occurs immediately to the left of the terminal. The guideline helps him with other Keys. For example, with its help, the learner can determine that the Key in *continuity* is *it*, not *uit*, and that the Key in *jealousy* is *ous*, not *us*. When the learner can find the Key accurately, he has the best chance of placing stress accurately; the guideline improves his chances of ultimately stressing a word correctly.

The guideline for the Key in strong rule words applies to all words with terminals. In addition to this guideline, the learner receives specific practice finding the Key in each new domain he studies. Each lesson also familiarizes the learner with the alternate forms the new terminals may have when suffixed. For example, the *-y* and *-acy* terminals change to *-i-* and *-aci-* before certain endings; the *-ate* and *-ize* terminals become simply *-at-* and *-iz-* before *-es*, *-ed*, *-ing* and *-er/-or*. These special pointers help the learner isolate the Key by helping him recognize the terminal.

Two-syllable nouns have come into Modern English from diverse sources and sometimes with archaic and foreign spellings. For this reason, the spelled composition of their Keys is unrestricted. The Key in short nouns is spelled with one or two vowel letters, e.g. *captain*, *ménu*.

*Predictive Accuracy.* Let us consider the success rate of the LSSR as applied to its two large domains, words with terminals, and short nouns. In the first nine lines of the list in (16), more than 8,000 words are represented. In the categories called 'Other Terminals,' there are approximately 2,000 words. The LSSR correctly predicts the major stress of these words no less than 95% of the time.

For short nouns, the story is different. Two-syllable nouns form the largest single category of words in the English language. Excluding those words which have weak endings, the remaining corpus of short nouns contains about 7,000 items. Of these, words having a V or VC in the Key are quite regularly stressed on the Left Syllable. However, words having a VCC, VCe or VV in the Key are stressed on the Key and on the Left Syllable with about equal frequency but without any phonological or morphological basis for stress placement.<sup>18</sup> I have made no complete count of short nouns stressed by the LSSR, so I can give no reliable estimate of stress assignment accuracy. My impression, however, is that the LSSR will generate correct stress for short nouns perhaps 75% of the time. The remaining words make up a large group of short nouns which the learner must memorize, as native speakers do, because their stress is unpredictable.

*Preparation.* To use the LSSR for words with terminals, the learner's preparation involves [1] becoming familiar with weak endings and strong sequences, [2] judging the number of spelled syllables a word has, and [3] locating the Key.

Requirement [1] is important for two reasons. First, the learner must not mistake a weak ending or a strong sequence at the end of a word for a terminal. Second, having found a terminal, the learner will need to identify any weak endings or strong sequences to the left of the terminal in order to determine which stress rule applies. To familiarize the learner with endings and sequences, the pronunciation course must introduce the LSSR after the other three rules. In this way, the learner will have the opportunity to work first with a variety of endings and sequences associated with the other rules.

Requirement [2] insures that the word which appears to have a terminal is indeed a long word. The learner's study of vowel spelling patterns enables him to enumerate spelled syllables. For requirement [3], the learner again needs his spelling pattern skills in order to separate terminals from the remainder of the word.

To use the LSSR for short nouns, the learner must be able to [1] identify nouns by context clues, [2] judge the number of spelled syllables in a word, and [3] locate the Key. These requirements are neither demanding nor new. His experience with the requirements comes mainly from other rules. However, his ability to count syllables is exercised in the identification of long words, as discussed above.

The two Strong Stress Rules discussed above provide the learner with the same prediction advantages that the Weak Stress Rules provide. The learner can assign stress to words without knowing the words beforehand but by using, in a mechanical way, information readily available to him. Then, having assigned the stress, he can predict the vowels in such words by combining his stress information with vowel spelling information. Thus, the four-rule stress system supplies the learner's needs while at the same time accommodating his limitations.

#### SPECIAL STRESS TOPICS

In his pronunciation course, the learner is introduced to stress topics in this typical way: Two to four word domains are taken up in a new lesson and the rules which govern their stress and vowel qualities are practiced in written and oral form.<sup>19</sup> Occasionally, and for different reasons, this pattern of presentation is broken in order to handle a special group of words. A few of these special topics are mentioned here to complete the picture of



the system which the learner uses to assign major stress.

One special topic is the stress of polysyllabic function words. This topic concerns words like *although, under, about, any*. While accommodated by the learner's rules, these words form a domain which is defined quite differently from any other and is therefore segregated for special treatment. Another topic is the stress of words with neutral endings (listed in note 4). Several lessons show the learner which stress rule applies to the remainder after removing the neutral ending. Since neutral endings are not properly stress domains, this topic has to be dealt with separately.

A third special topic is that of autostressed sequences, such as *-áde, -áire, -éé, -éer, -ése, -ésce, -ésque, -éur, -éque, -éte, -ier, -óo, -óon*. Words with these sequences are of foreign origin and deserve special consideration. Their stress pattern is the opposite of what is expected of words with terminals. Instead of carrying major stress on the third syllable from the end, these words ordinarily carry a minor stress in this position. The major stress is on the word-final autostressed sequence, the position where a terminal often has a minor stress.<sup>20</sup>

A fourth topic dealt with in a special way is the stress of short adjectives. This topic involves such a diverse, yet common, set of words that the learner must take care to use the applicable rules properly. A fifth topic--words ending in the single vowel letters *a, i* and *o*--require special handling, not because of their stress (VWSR), but because of their vowel qualities. Most of these words have been borrowed from Spanish or Italian, and as such, they preserve many evidences of their foreign origin in the quality of their stressed vowels.

The pedagogical stress system presented here accounts for the major stress of nearly 100,000 polysyllabic words. When monosyllabic words are added to this list, the learner can stress properly almost all of the vocabulary in educated English.

#### CONCLUSION

The bipartite pronunciation goals--*production/discrimination* and *prediction*--are solidly based on the nature of language as performance and competence (Dickerson 1980b). It is therefore a theoretical imperative to incorporate these goals into pronunciation instruction. But as a practical matter, is it possible for the learner to meet these goals? In particular, is the goal of

prediction realistic? Thanks to the system described above, the answer is, yes.

As noted at the beginning of this paper, the stress of a word is fundamental to its pronunciation. The stress defines not only the supra-segmental character of a word but also its segmental character, since it governs all vowel qualities and certain consonant choices. But more than this, word stress is at the heart of phrase rhythm, which ultimately and profoundly affects the communicability of an utterance. Because word stress is prerequisite to the prediction of so much in spoken English, it would be difficult if not impossible to implement fully the pronunciation goal of prediction without a means to predict word stress. A pedagogical word stress system, then, must be the keystone in any approach to pronunciation teaching which adopts as its own the two instructional goals motivated here.<sup>21</sup>

-----

FOOTNOTES

- <sup>1</sup>It is also my intention to chronicle in these notes the development of the pedagogical stress rules, in order to help the interested reader put into historical perspective the dozen or so papers which we have written on this topic since 1974.
- <sup>2</sup>Prator and Robinett (1972) say it nicely, 'Stress then is the key to the pronunciation of an English word, and the location of the accent should always be learned with the word...Unfortunately, there are no infallible rules for determining which syllable of a word should be stressed. Many times you will need to turn to the dictionary unless you hear the word spoken by someone familiar with it.' pp. 18-19.
- <sup>3</sup>ESL texts influenced in some measure by generative insights, however, attempt more than ever before to provide pointers about word stress placement. Although the pointers touch on only a minute part of the English lexicon, what is offered is better than what we find in taxonomic texts. Apart from our work, no comprehensive or unified stress-assignment system is yet available to the learner because such a system is deemed too esoteric.
- <sup>4</sup>Other neutral endings are *-kin, -less, -let, -like, -man, -ment, -ship, -some, -ward, -wise*. The neutral ending *-s* never has an *e* before it. If an *e* precedes a word-final *s*, the *e* must always be included with the *s* to make a weak ending.
- <sup>5</sup>The MSR is in fact a highly compact collection of subrules, some of which even have names, as we shall see.  
  
The terms, Key Syllable and Left Syllable, are mine and are not used by Chomsky and Halle. The terms nevertheless identify accurately the syllables affected by the technical stress rules.
- <sup>6</sup>Our research did not progress unflinchingly to the PWSR. A trail of successively closer approximations has been left in the literature. In Dickerson 1975a (written in 1974), 1975b, 1977c (written in 1975), and 1977d (written in early 1976), the RSR was reduced to a pedagogical form called the SISL Evaluation Principle. SISL is an acronym for 'Stress Initial, Strong or Left.' That is, stress the starting syllable (now the Key Syllable) if it is word initial; if it is not initial, stress it if it is strong (= strong cluster); if it is not strong, stress the syllable to the left of the starting syllable (now the Left Syllable). The SISL principle required the learner to judge strong and weak clusters by sound. In Dickerson 1975c, a first attempt was made to redefine syllable types in terms of spelling, because it was clearly unrealistic to expect the learner to know the sounds of novel words. No revision of the stress rule was proposed at this time. But in Dickerson 1977a (written in 1976), the rule ('Weak Prefix Stress Rule') was reformulated in two respects: First, the paper showed how spelling could be used instead of sound as a guide to clusters. Second, the study showed that spelling criteria were inadequate to block leftward movement of stress in all necessary cases. Therefore, the

revised rule also incorporates a prohibition against stressing a Left Syllable containing a prefix: If the Key is <V/VC>: Stress Left but not a Prefix; Otherwise: Stress Key. This rule appears in Dickerson 1977b (written in 1976) and in Dickerson and Finney 1977 under the name of Key Stress Rule. Then, in 1977, the rule appeared as the Weak Stress Rule in Dickerson and Finney 1978 and in Dickerson 1978 (both written in 1977). Experience with the new rule during 1977 and 1978 led to the conclusion that the formulation incorporating *both* blocking devices constrained the leftward movement of stress excessively. Finally, in 1978, we dropped the syllable evaluation from the rule entirely, leaving the rule as stated in (6), with the single blocking device. In addition, the term 'Otherwise' was replaced by the more explicit phrase, 'If you can't Stress Left.'

<sup>7</sup> In many cases an *i* may precede the weak ending, e.g., *-iable*, *-ial*, *-iar*, *-iary*, etc. Words having such sequences with *i* are stressed by the Key Strong Stress Rule for reasons discussed below.

<sup>8</sup> Said differently and more graphically, if you could crowd four columns of fifty words each onto a page, such an arrangement would give you a book more than one hundred pages long.

All word counts cited in this paper are given in round numbers. The counts are based on the word lists of Dolby and Resnikoff 1964 and Lehnert 1971.

<sup>9</sup> It will be noted that the *u* in these words sounds tense. By sound, it would appear to form a strong cluster. The *u*, however, in Chomsky and Halle's analysis is underlyingly lax and is tensed by rule (SPE 241 (23) III). By judging a cluster on the basis of spelling alone, it is possible to treat the *u* as the weak cluster it is.

<sup>10</sup> As stated in note 6, the VWSR preceded the PWSR in the developmental chronology. It was soon discovered, however, that the VWSR by itself worked too poorly for too many classes of words to be useful. The V and VC criteria were so ambiguous with respect to vowel tenseness that the prefix criterion had to be added in order to block the stressing of Left Syllables. Later, the V/VC part was dropped altogether for reasons mentioned in note 6. It was reluctantly resurrected in 1979 when it became obvious that it governed classes of words where the PWSR failed.

<sup>11</sup> Subject to further study, the 2,000-member group of nouns and adjectives ending in the single vowel letters *i*, *o*, and *a* may be included in the domains of the V/VC Rule. Except for certain well-defined environments, stress falls on the penultimate syllable: *the rotúnda*, *the confétti*, *a torpédo*.

<sup>12</sup> In all Dickerson papers from 1974 through 1978, the Strong Stress Rule (now the KSSR) applies not only to iV-endings but also to a set of iC-endings (*-ic*, *-ical*, (verbal) *-ish*, *-id*, *-ity*, *-ify*, *-ible*, *-igible*) and to some uV- and uC-endings (*-ual*, *-uary*, *-uent*, etc.; *-ula*, *-ular*). For stress purposes, these domains are fine; the Key to the left of these endings properly receives major stress. But when attention is given to the quality of the stressed vowel in the Key, two problems surface. First, the strong iV-ending words require one set of vowel quality statements and the strong iC-, uV-, and uC-endings require a different set. Second, for words with iC-, uV-, and uC-endings, the

quality of the stressed vowel in the Key Syllable is exactly that predicted in other classes of words for a stressed vowel in the Left Syllable. To reduce the number of vowel patterns associated with the Strong Stress Rule and to avoid the duplicate statements of vowel quality, the iC-, uV-, and uC-endings were eliminated as strong endings and incorporated into the domains of other stress rules in ways which put the stressed vowel in the Left Syllable. This streamlining in 1979 left the Strong Stress Rule with a single domain, Strong iV-sequences. Later that year, eV- and yV-sequences were added to the iV-sequences and the three domains were collectively called strong sequences.

13 Vowel quality evidence can be used to support decisions about stress-related word positions--Key and Left Syllables--because different vowel qualities are systematically associated with these positions. See Dickerson 1980a.

14 It should be noted that verbs and adjectives ending in *y* and words having *-ancy*, *-ency*, *-ary*, *-ery* and *-ory* endings have already been accounted for by the Weak Stress Rules. Words ending in *-acy* will be discussed below, under Application. However, words ending in *-ity* are included in the group of *-y* nouns treated here.

15 The Alternating Stress Rule has a long pre-generative history. The earliest reference I have found to the rule comes from an orthoepic guide printed in 1680. It reads, 'Accent in a Syllabl in Words of Many Syllabls, is most commonly used on a vowel in *the* Third Syllabl from the last.' *A Treatise of Stops, Points, or Pauses* 1680.

Chomsky and Halle's contribution was to convert this observation into a technical rule in which the exact domains of application are fully specified and to incorporate the rule into an internally consistent phonological system.

Our applied task has been to convert Chomsky and Halle's rule into a useable generalization for learners of English. As before, the conversion process began with a rule close to the technical rule. Then, as our pedagogical system emerged, the rule was gradually modified to fit into the new scheme.

16 Chomsky and Halle note this condition in their ASR (SPE 95-96).

17 Two general sets of words remain exceptional for the LSSR. Both involve morphologically complex words of Greek origin. In the first set are words with minor stress on the Key (in American English) before a *-y* terminal. Many of these words are stressed left of the Left Syllable. In this set are words ending in *-àrchy* (*híeràrchy*, *óligàrchy*), *-mòny* (*mátri-mòny*, *téstimòny*), and similar words like *épilépsy*, *mélanchòly*, *óρθοδόxy*. In the second set are words with three-syllable prefixes which require stress on the first of the three syllables. These prefixes form a small group: *cinema-*, *(en)cephalo-*, *entero-*, *helico-*, *hetero-*, *platino-*, *sidero-*. In words like *héterodox*, *síderoscope*, the stress is left of the Left Syllable.

18 John Ross comments on two-syllable nouns with tense vowels in the Key (Ross 1972). He says, 'As far as I can tell, it is impossible to predict whether stress retraction [movement of stress from the Key to the Left Syllable] will take place in disyllabic nouns (240)...It seems that

stress retraction is essentially random, and that whether or not a form is morphemically complex has few consequences for predicting its stress (241).'

A few Keys retain stress consistently, such as *-óon*, *-óo*, *-ése*, *-áde*. But apart from these cases, there is *no* general value in the statement often seen in ESL texts that in Latinate verb-noun pairs, major stress falls on the last syllable of the verb but on the first syllable of the noun (*to confli<sup>ct</sup>*, *a c<sup>on</sup>flict*; *to reb<sup>el</sup>*, *a r<sup>é</sup>bel*). There are just as many nouns which have the same stress as their verbal counterparts (*to rep<sup>o</sup>rt*, *a rep<sup>o</sup>rt*; *to contr<sup>o</sup>l*, *a contr<sup>o</sup>l*).

<sup>19</sup> During the first semester of his pronunciation course, the learner becomes acquainted with all four stress rules and one-third of their domains, as listed above. Recall that time is also given to learning vowel quality patterns and prefixes. The second semester of work shows the learner how he can extend his four rules to the remaining two-thirds of the domains (Dickerson 1981a, b).

<sup>20</sup> The KSSR is extended to accommodate these words: For Strong Sequences and Foreign Keys: Stress the Key Syllable.

<sup>21</sup> I wish to express special thanks to Rebecca Finney and Pearl Goodman for their helpful suggestions during the preparation of this paper.

---

## REFERENCES

- Chomsky, Noam, and Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper and Row.
- Dickerson, Wayne B. 1975a. Generating Enthusiasm with Generative Phonology. ERIC Documents ED 105 751.
- Dickerson, Wayne B. 1975b. Predicting Word Stress, Generative Rules in an ESL Context, *TESL Studies* 1:38-52.
- Dickerson, Wayne B. 1975c. The WH Question of Pronunciation: An Answer from Spelling and Generative Phonology. *TESOL Quarterly* 9:3.299-309.
- Dickerson, Wayne B. 1977a. Assigning Stress to Multi-Suffixed Words: Applications for TESL. *ITL, Review of Applied Linguistics* 36:71-88.
- Dickerson, Wayne B. 1977b. Generative Theory in TESL Practice. *The Modern Language Journal* 66:4.179-187.
- Dickerson, Wayne B. 1977c. Reflecting Linguistic Insight in Pedagogical Materials, The Case of English Phonology. *ITL, Review of Applied Linguistics* 33:27-43.
- Dickerson, Wayne B. 1977d. Review of Lionel Guierre's *Drills in English Stress-Patterns*. *TESOL Quarterly* 11:1.97-103.
- Dickerson, Wayne B. 1978. English Orthography: A Guide to Word Stress and Vowel Quality. *IRAL* 16:2.127-147.
- Dickerson, Wayne B. 1980a. Bisyllabic Laxing Rule: Vowel Prediction in Linguistics and Language Learning. *Language Learning* 30:2.317-329.
- Dickerson, Wayne B. 1980b. A Pedagogical Interpretation of Generative Phonology: I. Theoretical Foundations. *TESL Studies* 3:63-97.
- Dickerson, Wayne B. 1981a. *Learning English Pronunciation, Volume III, Word Stress and Vowel Quality, Part I*, unpublished.
- Dickerson, Wayne B. 1981b. *Learning English Pronunciation, Volume IV, Word Stress and Vowel Quality, Part II*, unpublished.
- Dickerson, Wayne B., and Rebecca H. Finney. 1977. A Generative Approach to Vowel Orthography. *TESL Studies* 2:88-110.
- Dickerson, Wayne B., and Rebecca H. Finney. 1978. Spelling in TESL: Stress Cues to Vowel Quality. *TESOL Quarterly* 12:2.163-175.
- Dolby, J. L., and H. L. Resnikoff. 1964. *The English Word Speculum*. Sunnyvale, CA. Lockheed Aircraft Corporation.
- Lehnert, Martin. 1971. *Reverse Dictionary of Present-Day English*. Leipzig: VEB Verlag Enzyklopädie.
- Prator, Clifford H., and Betty W. Robinett. 1972. *Manual of American English Pronunciation*. 3rd ed. New York: Holt, Rinehart and Winston.
- Ross, John Robert. 1972. A Reanalysis of English Word Stress, Part I. *Contributions to Generative Phonology*, ed. by M. K. Brame, 229-323. Austin: University of Texas Press.

-----





MODELS OF HUMAN MEMORY  
IN SECOND LANGUAGE ACQUISITION

Stephen B. Dunbar

This review explores the relationship between current models of long-term memory (LTM) and second language acquisition (SLA). LTM models from the fields of cognitive psychology and artificial intelligence are analyzed with attention to the mechanisms which store linguistic information and to the status of the linguistic unit in the storing process. SLA theories of Selinker and Krashen are shown to be inconsistent with these currently held views of LTM. In light of these inconsistencies, three criteria are proposed for a sensible model of SLA, namely, (1) that a model be semantically or conceptually based, (2) that it characterize forms of long-term organization and storage of L<sub>2</sub> information, and (3) that it be explicit and testable.

Psycholinguistic research in the areas of human memory and second language acquisition (SLA) has suffered from a lack of rigorous experimental investigation. In spite of the fact that large amounts of data have been collected during the past five to ten years, our present understanding of how these seemingly nebulous areas are related is extremely limited. Accordingly, many researchers in SLA have largely focused their attention on formulating abstract theoretical models which presumably explain, or, to use an expression that has trapped scholars of the discipline, "account for" the data collected in isolated experimental situations. While such models have effectively established SLA as a research discipline, they have not as yet created a strong empirical basis for accepting or rejecting alternative hypotheses in the field. Of particular concern here is that there has been no systematic attempt to examine aspects of human memory vis-à-vis second language (L<sub>2</sub>) learning in spite of the fact that many of the L<sub>2</sub> models currently embraced contain implicit assumptions about the structure of long-term memory (LTM).

This paper will review some recent models of semantic memory<sup>1</sup> in an effort to characterize their relationship to two currently held views on the second language learning process: (1) the interlanguage hypothesis (Selinker, 1972; Nemser, 1971), and (2) the monitor model of L<sub>2</sub> performance (Krashen, 1977). In the process, I will show that in spite of the empirical studies being generated by these hypotheses, our ability to explain the SLA process has not improved, in part because of a failure to come to

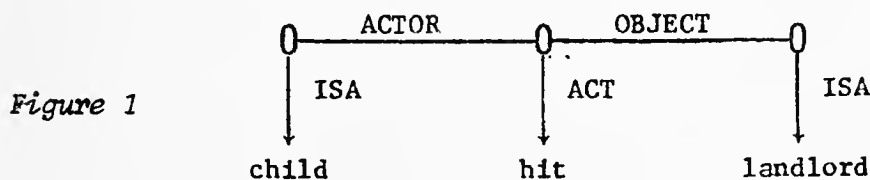
grips with the problem of memory in a second language situation. Specifically, SLA research has avoided direct contact with the question of psychological reality in language, even though this question is fundamental to any model of linguistic knowledge. As will be discussed later in the paper, the notion of psychological reality in L<sub>1</sub> studies is far from understood; nevertheless, the question is central when we are entertaining SLA models which postulate "latent psychological structures" (Selinker, 1972), or distinct "learned" and "acquired" systems (Krashen, 1977) operating during L<sub>2</sub> performance. Whether a more detailed investigation of what L<sub>2</sub> learners are able to remember would clarify this notion is problematic; however, such work would without question help in determining what happens in the mind of an L<sub>2</sub> learner as well as in providing an additional measure of the validity of current and future explanations of SLA.

Generally speaking, cognitive psychologists have defined human memory across two distinct theoretical constructs: an associative approach and a Gestalt or wholistic approach. The former, having been revised and expanded throughout its history (cf. Anderson and Bower, 1973), assumes that information is stored in memory by extensive links and associations with other chunks of information to which it is related. In its early form, associationism is well-known to language teachers as the theoretical basis for audiolingual methodology. Recently, associationism has been extended to the field of Artificial Intelligence, where it has been used to develop computer simulators of language use (Rieger, 1977; Winograd, 1971). Gestalt theories, in contrast, derive from experiments in the visual perception of ambiguous figures. They contend that memory for linguistic units, like that for visual forms, is intimately related to the notion of a structural unit in perception (Neisser, 1967). In other words, because of the nature of structural organization for a given stimulus, the memory trace for the whole unit (word, sentence, paragraph, etc.) is stronger than the sum of traces for individual parts (phonemes, syntactic categories, discourse units, etc.) of the stimulus. How these two basic paradigms have influenced models of memory of a linguistic unit, the sentence, is the focus of the following remarks.

The first model to be discussed in the associationist tradition is that of Rumelhart, Lindsey, and Norman (1972), who claim that LTM

for sentences is generated by a group of independent associations among the words of which a sentence is composed. The associations established in semantic memory are functions of various case relations (Fillmore, 1968) which obtain in particular sentences. Words themselves occupy nodes in LTM that are conjoined by relations such as actor-object, actor-act, act-object, etc. So for a sentence like "Fred ate dinner," associative links of actor-act (Fred ate) and act-object (ate dinner) would be encoded in LTM. This model further posits a token-type relationship for the words of a particular sentence and the case relations that link them (i.e., word combinations are instances of case relation types), implying that the truly stable elements in sentence memory are case relation types. They establish associations among the parts of the sentence.<sup>2</sup> Words are mere tokens, and as such, have no primacy of their own (Rumelhart et al., 1972, p. 216).

Figure 1, taken from Anderson and Bower (1972), is a schematic diagram of the memory representation for the sentence, "The child hit the landlord," and will serve to illustrate the Rumelhart, et al., version of the associative paradigm.



According to the Rumelhart et al., model, abstract concepts in this sentence have the names, 'child,' 'hit,' and 'landlord,' but are connected to actual lexical items only through a separate and independent chain of associations in LTM. This fact points to the defining characteristic of the strict association theory of memory: associations are not figural in any sense. That is, a cluster of concepts, say a combination, child-hit, is possible only if there is a pre-existing association between the individual concepts (Anderson and Bower, 1972, p. 595). Exactly how the presence of such an extensive system of associations might be tested is a question that has occupied a great deal of memory research, yet to which there has been no definitive answer to this point. What this portends for those interested in the L<sub>2</sub> learning process is this: if it is the case that independent associations govern LTM in some systematic way, and that language-related associations are but one part of a larger

memory structure, what is the exact nature of associations that link L<sub>2</sub> knowledge to the larger memory structure of the learner? To what degree are associations linked with (1) L<sub>1</sub> information of a similar type at levels of discourse, syntax and phonology, and (2) non-linguistic concepts similar to those of the Rumelhart, et al., model?

Anderson and Bower (1978) examined the autonomy of associations according to the Rumelhart, et al., model by testing the ability of their subjects to recall objects of sentences, given differential cues. They argued that given a pair of sentences with similar surface structures and identical objects, like: (1) The child hit the landlord, and (2) The minister praised the landlord, the probability that subjects recall 'landlord' should be greater for a cue which includes the subject of (1) and the verb of (2) than it should for other potential cues (i.e., a subject alone, a verb alone, and a subject and verb from the same sentence). The representation of associations for these and other similar pairs of sentences in LTM appears in Figure 2 (Anderson and Bower, 1972, p. 596).

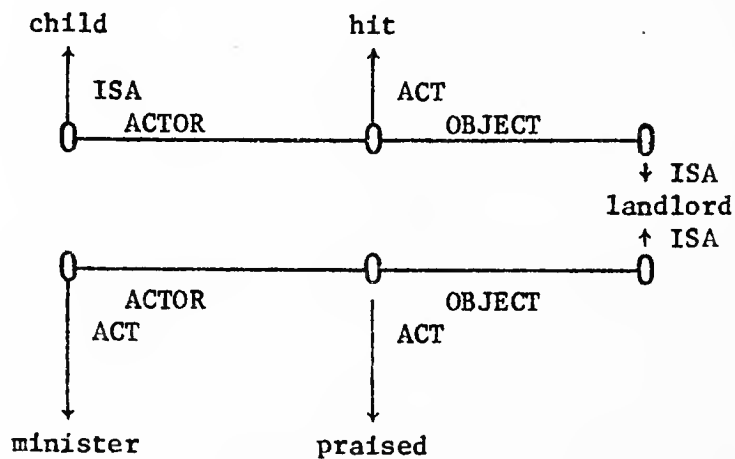


Figure 2

Anderson and Bower found the "cross-over cue," child-praised, to be the most effective in stimulating recall, suggesting that this type of cue opened two independent paths, or sequences, of associations leading to the same sentence completion, and hence was the optimal cue for recall. In this sense, a strict associationist view of LTM allows one to predict recall as a function of case relations that exist in an individual sentence, or between pairs of sentences.

The predictive power for certain cues in the retrieval of sentences from LTM has been the subject of other LTM experiments critical of the

strict associationism of the Rumelhart, et al., model, and of Anderson and Bower's own elaborate model of Human Associative Memory, or HAM (Anderson and Bower, 1973). Foss and Harwood (1975), in a replication of the Anderson and Bower experiments, found that object recall was significantly better when subjects were cued with a subject and verb from the same sentence. From this finding, they hypothesized that something about the organization of the combined SV cue, some kind of structural unity, perhaps, might have been producing better recall. In a variation on the Foss and Harwood theme, Anderson and Ortony (1975) also tested the primacy of associations over other possible variables in the LTM by using sentences of the following type:

Group I. The container held apples.

Group II. The container held cola.

The authors discovered that the semantic conceptions of the subject, which vary with the object of the sentence, were variables affecting recall. For example, a cue, "basket," initiated recall of a Group I sentence, whereas "bottle" facilitated Group II recall. Anderson and Ortony (1975, p. 172) claim this to be evidence showing that comprehension for, and memory of sentences involve the construction of elaborate and specific mental representations, and that a model based purely on the strings of independent associations doesn't offer a complete explanation for memory tasks that require more sophisticated semantic interpretations.

One observation that could be made at this point about the associative paradigm in memory theory is that it provides a precise, though somewhat overwhelming, characterization of the structure of semantic memory for sentences. Indeed, it taxes the imagination to conceive of such lengthy and elaborate association networks, let alone of an experimental procedure sufficiently sensitive to identify some of the exact associations that exist in LTM. This is clearly an issue that would have to be addressed in characterizing semantic memory in second language learners. Perhaps a more immediate question, however, is how might an associative system function in processing and remembering linguistic input, and whether a similar functioning process is operative for L<sub>2</sub> input.

Melvin (1977) and Melvin and Rivers (1976) have reported on contributions from the field of Artificial Intelligence to theories of semantic memory vis-à-vis foreign language teaching. The former correctly acknowledges

the difficulty in extrapolating concepts meaningful to the language teacher from computer simulators of language understanding (p. 99); however, models such as those presented by Winograd (1972), Rieger (1977), and Schank (1972) allow us to see in principle how an associative model of semantic memory actually "works" in the process of understanding linguistic input. While no systematic review of the literature on Artificial Intelligence is attempted here, the design of language simulators introduces another dimension in the relation between LTM and L<sub>2</sub> learning models.<sup>3</sup>

Winograd (1972) implements what he calls a procedural approach to language comprehension. He interprets language use as calling on all of human knowledge and intelligence, and his language simulator represents such knowledge as "procedures," operations that can be carried out by a computer to generate meaning from syntactic groups in a sentence. These procedures work in conjunction with (1) a parser, which identifies and separates syntactic groups (noun groups, verb groups, etc., according to Halliday's (1970) systemic grammar); (2) a recognition grammar of English; (3) programs for semantic analysis; and (4) a problem-solving system. The operation of Winograd's model is restricted to an artificial domain wherein a robot on an output screen manipulates physical objects at the command of the computer programmer and maintains a dialogue while so doing. The robot is able to carry out commands, ask for clarification about, for example, unfamiliar vocabulary, retrieve knowledge from past experience, and use it to respond to new situations it encounters within the artificial world. Keyboard input to this model in the form of an English sentence invokes a series of procedures wherein noun and verb groups are isolated and subjected to semantic analysis; inferences are drawn about the most plausible interpretation of the sentence and the appropriate action and response are output. While Winograd's notion of the procedure is at the core of his model's activity and is thus basic to LTM for linguistic information, another feature obtrudes. In its capacity to maintain a dialogue during operation, the model must be sensitive to language at the level of discourse (cf. Winograd, 1970, pp. 32-33), and must accordingly store such knowledge in addition to sentence level syntax.

Schank (1972) and Rieger (1977) adopt a radically different attitude toward language simulators, each arguing that language is conceptually based (cf. Schank, 1972, p. 553). Accordingly, they invoke such terms as conceptual dependency (Schank, 1972) and word sense network (Rieger, 1977) in explicating essential features of their models. Each argues for a language understanding device which, as Rieger (1977) claims, should be "a concise framework for allowing language to sprawl throughout a larger system of intelligence" (p. 97). At the heart of Rieger's theory is the assumption that language in LTM is organized in terms of word sense networks, each being defined as "an aggregation of expertise about one word of a language" (p. 102), and not in terms of syntactic categories, grammatical relations, and the like. Syntax in this case is understood to be a time-saving device which makes accurate selection of the appropriate word sense in a given sentence more efficient. In the program itself, a syntactic parser analogous to that of Winograd is operative, but in a different way. Rieger's parser searches for the main verb of the sentence; once this is found, a "sense selection network" interprets the word sense of the main verb. This word sense acts as a guide for subsequent syntactic parsing. According to Rieger, this approach allows for the primacy of meaning -- some core of concepts -- in language comprehension.<sup>4</sup>

What the models of Winograd and Rieger reveal relative to semantic memory for sentences is that a workable model of language understanding requires the storage of information from all levels of language, but in particular, the syntax and the lexicon. Presumably, connections between these levels of processing are important variables in the successful functioning of language simulators -- in principle, we might conjecture that similar connections are necessary for successful understanding in a second language, and that storage of word senses, for example, is fundamental to language learning. Rieger himself indicates that "the things that concern a child in the earliest phases of language are the times (contexts) at which it is appropriate to utter a word" (pp. 125-6). Language development, in his view, consists of subsequent refinement of word sense networks and the constant addition of new ones.

Rieger's remarks on language learning may or may not be true; nevertheless, the discussion and criticisms of associative models of LTM and

language simulators following their lead show much disagreement about the exact form of storage in sentence memory. All models so far mentioned support the assumption that language is organized -- the question plaguing most applied linguists, psychologists and computer scientists is "in terms of what?" The first section of this paper has examined the mechanism, the association, for organizing whatever it is language users store when they process sentences. In what follows, research into the form (linguistic or otherwise) of stored language information is reviewed in an effort to discover the status of the linguistic unit, since this problem must be accommodated in any comprehensive theory of SLA.

Neisser (1967) proposes that Chomsky's investigation of language as rule-governed behavior, and of the sentence as an organized hierarchy of constituent structures, has created a union of the generative grammarians of the early 1960's and the Gestalt psychologists of the earlier part of the century (pp. 243 ff.). Specifically, such insights into the syntactic structure of language have led to various hypotheses about the linguistic nature of memory codes. Underlying all such hypotheses, however, is the belief that linguistic rules of some sort, as principles of organization, are psychologically real. This notion goes right to the heart of many analyses of L<sub>2</sub> learning, the contrastive analysis and interlanguage hypotheses most notably, but to the monitor model as well in that Krashen's "learned" system in part consists of rules acquired during formal language instruction. The general parallel between rules and memory constructs lies in the notion of figural constancy being operative in sentence processing (perhaps coded in syntactic relations based on transformations). Presumably it is this property of sentences, and of language itself, which allows for long-term storage.

The search for psychological correlates for these linguistic structures first took form in what Fodor, Bever, and Garrett (1974) call the coding hypothesis. Roughly characterized, the coding hypothesis states that the stored representation of a sentence is somehow identified with its linguistic base structure. Although from this point of view memory stores such material independently of the transformations which account for the sentence's surface form, both base struc-



tures and transformations are represented in LTM according to the coding hypothesis.

Mehler (1963), in an early experiment testing the coding hypothesis, studied the effect of grammatical transformations on free and prompted recall tasks. He found the number of transformations needed to derive a sentence syntactically to be inversely proportional to its frequency of recall; in other words, the fewer the transformations, the better the recall. Thus it appeared to be the case, as the coding hypothesis would predict, that recall for lexical content was independent from recall of passive, negative, WH-question, or other syntactic types. Recall of lexical items, part of the base structure, showed differential success and was therefore interpreted as support for the coding hypothesis.

Bever and Mehler (1967) found additional support for the coding hypothesis when they compared memory for sentences containing sentential adverbs to those with verb phrase adverbs, as in the following pair:

- (1) John probably ate the meat-loaf sandwich.
- (2) John reluctantly ate the meat-loaf sandwich.

Since the location of these elements in deep structure distinguishes syntactic nodes, that is, sentence versus verb phrase, it was reasoned that differences in recall would reveal something about the way these sentences are coded in memory. The results showed improved recall of sentences with sentential adverbs when the adverb appeared in initial position, and better recall of sentences with verb phrase adverbs when the adverb was contiguous with the verb. Errors in recall indicated a general tendency for subjects to locate adverbs in their position in deep structure.

The goal of the above experiments, and of the coding hypothesis as a psychological theory of language, is the development of an explicit deep structure account of linguistic entities as psychologically real constructs. Seen in the context of previously mentioned approaches to the representation of linguistic knowledge in memory, the coding hypothesis is but one linguistic explanation. It has also been argued (cf. Fodor, et al., 1974) that surface structure variables influence recall for sentences. The salience of an adverb at the beginning of a sentence in the Bever and Mehler experiment might be interpreted as an example of this phenomenon. Foder, et al., (1974) report on a group of studies designed

to support this view, known in general terms as the mean-depth hypothesis, which establishes the number of left branches contained in a sentence as an index of difficulty in recall. In other words, a sentence with more embedded constituents is more difficult to remember than one with fewer embeddings. Although proponents of this view argue that this notion of depth is transparent in surface structure, Fodor et al., (1974) conclude that in spite of the fact that subjects do use surface information in recall, "currently available data provide no convincing evidence that any syntactic description of a sentence is isomorphic to its representation in long-term memory" (p. 270).

This judgment of Fodor, Bever, and Garrett finds support in a host of studies showing that people remember meaning more frequently than they do syntactic form (cf. Fillenbaum, 1966; Sachs, 1967; Brewer, 1975). Fillenbaum (1966) began with the hypothesis that memory for gist is better than verbatim memory, and designed an experiment to examine recall for sentences containing antonym pairs in various combinations and logical sequences. The variables pertinent to choice of the antonym pairs were (a) morphological contraries -- words with un- or dis- prefixes, (b) logical contraries indicating a scalar property like hot/cold, and (c) logical contradictories such as open/closed. Fillenbaum's subjects were presented with 96 randomly sequenced, unrelated sentences containing combinations of antonym pairs of these three types. Sentences were of the following form:

A is X., B is not X., C is  $\bar{X}$ ., D is not  $\bar{X}$ ., where X and  $\bar{X}$  are an antonym pair. Results revealed significantly more meaning-preserving errors in recall than meaning-changing ones.

Fillenbaum's data conflict with the central assumption of the coding hypothesis, that memory stores sentences in terms of base structures and transformational rules. For example, rather than just "dropping out" entirely, negatives shifted to an antonym such that meaning was preserved; B is not X., thus became B is  $\bar{X}$ . Moreover, a large number of meaning-preserving errors was also discovered when (a) antonyms with the same root were available (e.g., pleasant vs. unpleasant), and (b), when items were phonetically distinct but were at the same time logical contradictories (e.g., wet vs. dry). Fillenbaum concludes that not only do subjects preserve meaning during recall, they also use

their knowledge of contradictory and contrary relationships in performing memory tasks. He argues that "in coding sentences for storage...subjects must respond discriminatively to features represented or derived from the conceptual distinction between these different sorts of opposites." (p. 227).

Brewer (1975) takes this notion of memory for gist one step further in a series of synonym-substitution experiments. He had subjects memorize random lists of sentences containing words for which there are potential synonym substitutes. Items were controlled for factors like naturalness of a substitution and degree of equivalence in meaning. These judgments were made by the subjects themselves. Like Fillenbaum, Brewer found that sentence recall maintained meaning better than it did surface structure. In fact, two of his experiments showed that subjects were more likely to produce a new list of equivalent items (shifted synonyms) than a list of the presented forms.

Theories of memory based either on surface structure or underlying syntactic representations cannot account for Brewer's data; the former because they lack an abstract level of analysis for words in memory, and the latter because they do not represent lexical items in memory, only base structures and transformations. Brewer also rejects the explanation that sentences are stored as mental images, since this requires a difference in recall between concrete and abstract sentences, which he found in another experiment to be false (Brewer, 1975).<sup>5</sup> Almost by default, Brewer introduces the notion of memory-for-ideas, a model in which sentences are coded in "non-linguistic, non-image, abstract representations," and are remembered through "underlying ideas, with parallel storage of some image information and some retained surface information" (p. 464).

At first glance, Brewer's position seems an elaborate and carefully-reasoned hedge; the memory-for-ideas hypothesis simply states that many factors are involved in semantic memory, and that no one-dimensional theory accounts for all language-related phenomena that have been observed in LTM research. On the other hand, Brewer's analysis may well be the most accurate characterization one could hope for, given the empirical complexity of studying interactions between human language and human cognition. In any event, his work serves to illustrate something underlying this entire discussion of LTM, namely, that there is much disagreement as to

as to the status of the linguistic unit which is actually stored in memory for sentences. The associative models were very explicit as to the mechanisms involved in the storage of information, but not consistent about the autonomy of associations and the extent to which discrete units or wholistic ones were operative. Computer simulation devices were in disagreement about the roles of syntax and lexis in working models of human memory. Studies identifying the linguistic form of sentential information in LTM indicate that meaning has primacy over structural description, but that neither is alone sufficient in explaining recall. Thus, Brewer's contention that memory has some global characteristics has some appeal, especially when we turn to the realm of second language acquisition, where we observe a complex matrix of linguistic, cognitive, and cultural variables interacting in a way that confounds nearly every effort to explain the L<sub>2</sub> learning process.

What the theories of memory reviewed here do have in common is the notion that LTM has an organizing function; it is an understanding of this property which Brown (1972) feels is crucial to the formulation of a coherent theory of second language acquisition. Most human learning, Brown argues, depends on "an efficient conceptualizing process of organization" which is basic to retaining meaningful information in LTM. With this *a priori* relationship between learning and memory, we might ask a more immediate question: to what extent do the various models of organization in memory lend support to current theories of second language acquisition?

Selinker's (1972) notion that L<sub>2</sub> learning is controlled by a latent psychological structure in the brain parallels Lenneberg's account of a child's ability to acquire his/her native language. The interlanguage hypothesis, in defining fossilized forms as "linguistic items, rules, and subsystems which speakers of a particular NL will tend to keep in their IL relative to a particular TL," finds its greatest support from theories of memory which posit linguistic rules as psychologically real constructs. From this point of view, an L<sub>2</sub> learner's performance may be described through attributes of memory which bear directly on linguistic entities. However, the comments of Fodor, et al., (1974) on the coding hypothesis should signal a warning about describing phenomena of language and memory on purely linguistic grounds. In particular, we

might wish to expand Selinker's definition of the psychologically relevant data in L<sub>2</sub> acquisition such that it included, for example, cultural aspects of concept formation, consistent with the notion that language is at some point conceptual in nature. Rieger (1977) provides evidence that not only is language, and presumably interlanguage, conceptual, but that this conceptual base is also, at the level of word senses, intimately related to the linguistic unit itself. Hence, if we view L<sub>2</sub> learning as a process drawing on all aspects of a learner's psychological identity, we would need to characterize the influence of general cognitive functions like concept formation on a learner's LTM when placed in an L<sub>2</sub> situation.

Like the interlanguage hypothesis, Krashen's monitor model for second language performance posits the existence of separate linguistic, and presumably cognitive, systems to describe idiosyncratic and systematic errors in L<sub>2</sub> performance. In particular, Krashen (1977) argues for a theoretical distinction between the knowledge acquired from informal or unconscious learning situations, and that learned through conscious efforts, in an L<sub>2</sub> classroom, for example. This latter capacity is predicated on the learner's having reached Piaget's period of formal operations in cognitive development. McLaughlin (1978) has criticized the monitor model on the basis of its suspect separation of conscious and unconscious learning. From the point of view of LTM research, it is similarly suspect to dichotomize what LTM stores as exclusively acquired or learned. The associative theories of Anderson and Bower, Winograd, and others do claim that links between memory nodes are autonomous, and that nodes themselves are discrete; however, nowhere in these models is there a precedent for completely separating knowledge or experience solely in terms of the context in which it enters the mind. On the other hand, Gestalt approaches to semantic memory indicate that perception of similarities (or, to borrow McLaughlin's terms, structural unity in what we "feel" and what we know by "rule") in the variety of experience one has with an L<sub>2</sub>, would in some way be reflected in the LTM of L<sub>2</sub> learners.

From another point of view, the monitor model fails when we examine Brown's (1972) notion of cognitive pruning as a basic psychological heuristic used by learners to encode material in LTM. Regardless of when and where an L<sub>2</sub> learner attends to the language he is learning, his tendency is to identify larger concepts into which smaller bits of knowledge can be placed. These "subsumers" would imply that a learner's natural

strategy for remembering L<sub>2</sub> input is to look for similarities between his past and present experience with the L<sub>2</sub>, since this would facilitate recall. To say that such experience is encoded as a learned system and an acquired system is tantamount to arguing for the arbitrary invocation of some new kind of memory structure unique to L<sub>2</sub> performance situations, something that functions in a cognitive vacuum whenever time and circumstances are appropriate.

These considerations aside, we might then ask what specific demands LTM research places on a model of the L<sub>2</sub> learning process if the models currently available are inadequate, and what kind of empirical research is needed to support a model that would meet these demands? While it is not my purpose in this paper to present anything like a theory of language learning, some considerations as to the criteria such a model should meet are in order, given the characteristics of long-term storage presented earlier.

First of all, both associative and Gestalt theories of LTM indicate that any reasonable model of L<sub>2</sub> learning should be semantically, perhaps even conceptually, based. Note that this is not an argument against the axiom that language learning entails (along with other things) the mapping of a new set of syntactic and phonological features onto an already existing semantic network. Indeed, it is a characteristic of nearly every theory of memory for language that some universal, language-independent entity is at the core of our organization of knowledge. However, the way in which this memory structure works in an L<sub>2</sub> situation is far from clear, and is a question which has received very little attention among the growing number of applied linguists studying second language acquisition.

Second, and perhaps a corollary to the first, a model for L<sub>2</sub> learning should characterize how L<sub>2</sub> information is organized and stored by L<sub>2</sub> learners. Clearly, learner strategies are closely related to this question, as are what McLaughlin (1978) refers to as schemata and discovery procedures. What we don't know as yet is whether such schemata, or organizing procedures, are constant, or whether they change as the learner's experience with the L<sub>2</sub> in various situations increases. As McLaughlin (1978, p. 321) also notes, these cognitive structures may exist on many levels of language-use -- semantic, syntactic, phonological,

pragmatic -- providing even stronger justification to attempting empirical examinations of them prior to developing any reasonable theoretical model.

A third criterion for a theory of language learning is that it should be explicit and testable. One approach to this is the simulation of an L<sub>2</sub> learner by a computer.<sup>6</sup> While I would hesitate to offer this as the litmus test for an acceptable model of L<sub>2</sub> learning, it would compel researchers to be explicit about the process they investigate. A less artificial means of attaining explicitness would be to allow our theories to grow out of empirical facts rather than to squeeze the facts from theories.

Empirical support for any strong claims about LTM for material in a second language is severely lacking. Lado (1965) reports a number of studies related to the immediate memory span of a group of ESL students in the United States. Glöcksberg (1963), for example, found that immediate memory in his subjects was sensitive to the meaningfulness of the stimulus material presented (serial digit items versus items in a linguistic context). Immediate memory represented materials in terms of 'chunks.' His data also showed that in the five weeks during which his subjects were enrolled in an intensive English program, memory span increased significantly for material presented in context as opposed to the serial digit items.

Unfortunately, the studies of memory span reported by Lado do not reveal any explicit information about the representational form of stored data. In contrast, Henning (1973) investigated the importance of acoustic and semantic factors in recognition memory for L<sub>2</sub> vocabulary, and found that language learners encode vocabulary in LTM both in terms of acoustic and semantic variables. In addition, he observed a developmental sequence which correlated with language proficiency. Just as younger children tend to encode acoustic signals in memory for words, so L<sub>2</sub> learners at lower levels of proficiency produced acoustic errors on Henning's recognition task. Subjects of higher levels of proficiency tended to use more semantic characteristics in recognizing L<sub>2</sub> vocabulary.

It is obvious that Henning's data do not offer a comprehensive picture of an L<sub>2</sub> as conceived in memory. They do, on the other hand, provide concrete information about variables related to the representation of L<sub>2</sub> input in LTM and in this sense, they point to a new direction in

SLA research. Were studies of this nature expanded to, in particular, the level of discourse, we would be better able to describe how L<sub>2</sub> learners do, in fact, integrate lower-level acoustic and syntactic processing with the ability to remember meaning. This fourth criterion, that a model characterize a learner's competence above the level of sentence processing, is no less crucial than the others to an overall picture of SLA.

What then can be said of the relationship between models of human memory and SLA? On the basis of evidence reviewed above, empirical foundations for describing this relationship are at present very unsteady. Extrapolation from general studies in memory for language is always possible, but a better approach is to confront the problem of memory directly. Understanding this aspect of SLA is, therefore, an empirical endeavor; whether it will be approached as one is a question for the future.

-----



NOTES

- <sup>1</sup>The terms, long-term memory, and semantic memory, are used interchangeably throughout this paper.
- <sup>2</sup>Anderson and Bower (1973, p. 207) also describe this distinction in terms of "idea" nodes and "word" nodes. The former refers to an abstract, non-linguistic concept, and the latter a lexical item which is, in turn, linked with orthographic and phonetic information. One idea node may be connected to more than one word node, and vice-versa.
- <sup>3</sup>One should observe that language simulators as such are not models of language learning. Indeed, Anderson and Bower (1972, p. 131) make a special note of there being no theory of language acquisition sufficiently explicit to be programmed on a computer. This problem aside, Rieger (1977) argues that his model of language comprehension actually reflects the way in which language is learned at early stages by children.
- <sup>4</sup>An important note of contrast here is that Winograd's program, though restricted in context, does what it was designed to do, responde to commands and maintain a dialogue. Rieger's model, on the other hand, is only in its formative stages -- the task of programming a machine to identify word senses is formidable.
- <sup>5</sup>The notion of LTM as an organized set of mental images is yet another theoretical viewpoint on semantic memory and beyond the scope of this paper. For a detailed account, see Paivio (1971).
- <sup>6</sup>Pioneer work has begun in this area, though I am not sure with what success. Scarcella and Baker (1978) report briefly on a paper of Hans Lee and Paul Munsell entitled "Computer simulation models of vocabulary learning in second language acquisition." See *SLANT*, 1978, 8 (2):64.

-----

REFERENCES

- Anderson, John R., and Gordon H. Bower. Configural properties in sentence memory. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11(5):594-605.
- Anderson, John R., and Gordon H. Bower. *Human Associative Memory*. New York: 1973.
- Anderson, Richard C., and Andrew Ortony. On putting apples into bottles: a problem of polysemy. *Cognitive Psychology*, 1975, 7(2):167-180.
- Bever, T. G., and J. Mehler. The coding hypothesis and short-term memory. *AF Technical Report*. Cambridge, Mass.: Harvard Center for Cognitive Studies, 1967.
- Brewer, William F. Memory for ideas: synonym substitution. *Memory and Cognition*, 1975, 3(4):545-548.
- Brown, H. Douglas. Cognitive pruning and second language acquisition. *Modern Language Journal*, 1972, 56(4):218-222.
- Chomsky, Noam. *Language and Mind*. New York: Harcourt Brace Jovanovich, Inc., 1968.
- Fillenbaum, S. Memory for gist: some relevant variables. *Language and Speech*, 1966, 9:217-227.
- Fillmore, Charles J. The case for case, in E. Bach and R. Harms, eds., *Universals of Linguistic Theory*. New York: Holt Rinehart and Winston, 1968:1-88.
- Fodor, J., and T. G. Bever. The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior*, 1965, 4(4):414-420.
- Fodor, J., T. G. Bever, and M. Garrett. *The Psychology of Language: An Introduction to Psycholinguistics and Generative Grammar*. New York: McGraw-Hill, 1974.
- Foss, Donald J., and David A. Harwood. Memory for sentences: implications for human associative memory. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14(1):1-16.
- Glicksberg, Daniel H. A study of the span of immediate memory among adult students of English as a foreign language. Unpublished doctoral dissertation, University of Michigan, 1963.
- Halliday, M. A. K. Notes on transitivity and theme. *English Journal of Linguistics*, 1967, 3:37-81.
- Henning, Grant H. Remembering foreign language vocabulary: acoustic and semantic parameters. *Language Learning*, 1973, 23:185-196.
- Krashen, Stephen. Second language acquisition, in William Orr Dingwall, ed., *A Survey of Linguistic Science*, 2nd ed. Stamford, Conn.: Greylock, Inc., 1977. pp. 317-338.
- Lado, Robert. Second language memory span. *IRAL*, 1965, 3(2):123-129.
- McLaughlin, Barry: The monitor model: some methodological considerations. *Language Learning*, 28(2):309-332.

REFERENCES (continued)

- Mehler, J. Some effects of grammatical transformations on the recall of English sentences. *Journal of Verbal Learning and Verbal Behavior*, 1963, 2:346-351.
- Mehler, J., and P. Carey. The role of surface and base structure in the perception of sentences. *Journal of Verbal Learning and Verbal Behavior*, 1967, 7:335-338.
- Melvin, Bernice. Recent developments in memory research and their implications for foreign language teaching, in J R. Cowan, ed., *Language for Special Purposes*. Urbana, Illinois: Unit for Foreign Study and Research, 1977. 89-110.
- Melvin, Bernice, and Wilga Rivers. In one ear and out the other: implications of memory studies for language learning. In J. Fanselow and R. Crymes, eds., *On TESOL '76*. pp. 155-164.
- Neisser, Ulrich. *Cognitive Psychology*. New York: Appleton-Century-Crofts, 1967.
- Nemser, William. Approximative systems of foreign language learners. *IRAL*, 1971, 9(2):115-123.
- Rieger, Charles J. Computational linguistics, in William Orr Dingwall, ed., *A Survey of Linguistic Science*. 2nd ed. Stamford, Conn.: Greylock, Inc., 1978. pp. 97-134.
- Schank, R. Conceptual dependency: a theory of natural-language understanding. *Cognitive Psychology*, 1972, 3(4):552-631.
- Selinker, Larry. Interlanguage. *IRAL*, 1972. 10(3):209-231.
- Winograd, Terry. Understanding natural language. *Cognitive Psychology*, 1972, 3(1):1-191.

-----



MICRO-ESL: A METHOD FOR TEACHING  
THE FUNCTIONAL/NOTIONAL SYLLABUS  
AND DEVELOPING COMMUNICATIVE COMPETENCE

Joan E. Friedenberq  
Curtis H. Bradley

Developing communicative competence is a major goal in the teaching of English as a second language. The functional/notional syllabus was developed in response to this goal, but few structured and systematic classroom techniques have been developed to implement it.

Microcounseling is one of the most widely used and systematic methods of developing interpersonal communication skills. By modifying this approach to include lexical, syntactic, and cultural components, we too, can use this effective method to teach ESL students to communicate effectively.

Developing communicative competence is a major goal in the teaching of English as a second language. The functional/notional syllabus was developed in response to this goal. Progress has been made in the development of effective functional content, but unfortunately, little has been done to develop a sound methodology to help the ESL teacher implement the teaching of this content. The purpose of this paper is to introduce micro-ESL, which has recently proved to be an effective tool in implementing the functional/notional syllabus and in developing communicative competence.

COMMUNICATIVE COMPETENCE AND THE FUNCTIONAL/NOTIONAL SYLLABUS

Since the late 1960's and early 1970's, language teachers have taken a greater interest in teaching students to communicate, as opposed to merely training them to attain competence. One pioneer in the area of communicative competence was Savignon (1972), who compared the communicative competence of three groups of college students learning French. The first group followed a traditional "modified ALM approach," and used the language lab once a week. The second group followed the same traditional approach, but instead of using the lab, they engaged in French culture activities (e.g., movies, foods, etc.) one time a week. The third group also followed the modified audio-lingual method, but they engaged in communication activities with native speakers once a week. At the end of the course, all students were given a traditional discrete-point exam in addition to a test of oral communication. Though there were no significant differences among group performance on the discrete-point test of linguistic competence, the third group performed considerably better on the test for communicative competence.

The results of this study demonstrate that students learn to do what they have opportunities to do. That is, if students are not given the opportunity to integrate the language components we so carefully teach them, they will not learn to use the language for communication purposes.

The communication activities which Savignon designed for her third group consisted of: 1) discussion, 2) information getting and giving, 3) reporting, and 4) description. Students were placed into groups of ten, and one native or near-native speaker of French joined each group. Activities such as information getting and giving, reporting, and description usually took place between one student and one native speaker. The two then exchanged conversation for one or two minutes in a simulated situation. Discussions were carried out with an entire group (one native speaker plus ten students) and topics related to current events. The situations and topics were generally defined in English (the students' native language). Savignon strictly imposed a philosophy of no grammatical corrections during communication times and the right to resort to the native language when necessary. Real communication was clearly the goal. Like Savignon, Chastain (1971:316) states,

To learn to use a language, the students must reach a point at which they can concentrate on what they are saying instead of how they are saying it...

Only a few have since attempted to develop classroom techniques for developing communicative competence. Farid (1976:300), for example, examined communication in the language classroom and identified five criteria by which a teacher can develop a good communicative activity.

A dialogue activity should be structured so that:

- a) the topic is interesting;
- b) the topic does not lie outside their [the learners'] semantic skills;
- c) the students engaged in the dialogue participate more or less equally;
- d) the participating students experience a feeling of success, regardless of the correctness of their English;
- e) the non-participating students are motivated to listen to content rather than to form.

Farid suggests that students practice their discussion in pairs and then

present them in front of the class.

Taylor and Wolfson (1978) have pointed out that although traditional language teaching approaches seem to be less effective than functional approaches, conversation in the language classroom must not be free and unstructured. They suggest that language teachers construct speech situations and provide students with the appropriate vocabulary, syntax, and socio-linguistic rules for the chosen situation. Like Farid's technique, students are then to practice conversing in couples.

Although there have recently been several valuable contributions in the area of communicative or functional/notional curriculum and materials development (Wilkins 1976, Stratton 1977, Fox 1978, Sheraga 1980, and Findley and Nathan 1980), little has been done to adapt a structured and systematic teaching methodology to this kind of syllabus. This concern has been recognized by Corder (1973), who expressed apprehension over whether communicative competence could be taught systematically in the classroom, and by Johnson (1977:673), who says:

It seems likely that simulation and role-playing will prove fruitful techniques, but their use in language teaching is as yet relatively new, and it will be several years before the materials developer has stocked a sufficiently large and varied armoury of such techniques to make a large-scale production of adequate functional materials feasible.

The technology of a somewhat modified microcounseling approach may provide the functional/notional syllabus with a structured and systematic technique that will make possible the development of both linguistic and effective interpersonal communication skills for adult ESL students.

#### MICROCOUNSELING

Microcounseling is a form of microtraining. The generic term, microtraining, refers to a general training format which is characterized by the development of specific, concrete skills through observation, practice, and feedback in a psychologically safe learning environment. Positive supervision is also an essential component of microtraining. Microtraining exists in a number of forms such as microteaching (Allen 1967), which is used to help pre- and in-service teachers-in-training develop specific teaching skills. Microsupervision (Chase, Coty, and Cotrel 1971) uses the microtraining format to teach supervisory conference skills. Microcoordination (Harrington 1970) uses

the microtraining format to teach job placement skills to cooperative vocational education coordinators. Of interest to the ESL teacher is microcounseling (Ivey, et al., 1968), which extends the microtraining format to the development of a large variety of effective interpersonal skills.

Microcounseling evolved from attempts to de-mystify the counselor education process. Counselor educators had for years been attempting to teach counselor trainees essential but elusive concepts such as "warmth" and "empathy." Success was limited because although everyone "knew" what was meant by these terms, it was difficult to define them operationally. That is, it was easy to identify a counselor who was being warm and empathetic, but difficult to identify exactly why. It was therefore exceedingly difficult to help trainees learn and use these elusive concepts. Ivey and his colleagues (1968) applied a component-skills approach to the interviewing process. This behavioral analysis of one of the important aspects of counseling resulted in the identification and definition in performance terms of a number of discrete behaviors which are component skills of effective interpersonal communication. From this initial research, a conceptual framework and technology evolved that has enabled Ivey and others to extend microcounseling far beyond counselor education and to behaviorally define other useful interpersonal skills.

Other applications of microcounseling include using it to improve the interpersonal skills of psychiatric nursing personnel (Hearn 1976), medical students (Authier and Gustafson 1974), and as media therapy with hospitalized psychiatric patients (Ivey 1973). Aldridge and Ivey (1975) demonstrated that junior high school students could be taught specific microcounseling skills as easily as adults. Bradley (1977) used microcounseling training as a direct, systematic interpersonal skills development program for inner-city youth. A variety of other applications can be found in Ivey and Authier (1978).

Although originally designed for trainers to work on a one-to-one basis with trainees, this method proved to be impractical for those with large number of students. Consequently, a second approach was developed which treats students/trainees on a group-wide basis.



### The Group Microcounseling Format

There are five parts to a group microcounseling exercise. The topic for the following exercise is "The Job Interview."

1. Introduction by the teacher. The students are informed that they are working on how to participate in a job interview.

An important part of a successful interview is listening to the interviewer and conveying to the interviewer that you are listening. Students are told that only that aspect of interpersonal communication should be considered, and not to worry about other dimensions for the time being. Their only goal then is to convey to someone that they are listening.

2. Training. The teacher asks for volunteers to role-play an employment interview. The job applicant is to "ham it up" and do everything wrong in terms of listening as is possible. By requiring the student to act as inappropriately as possible, she/he is placed in a no-lose situation. That is, if the student conveys very well that she/he is not listening, the student will be successful in terms of the exercise. If, on the other hand, the student does not convey this well, it will only show that the student is too good a listener. In addition, a good deal of humor is injected into the exercise by having students first act inappropriately. The interviewer is to play it "straight" and be as business-like as possible. The class is divided into groups of six each. They observe the role-played interaction (3 to 5 minutes is suitable for making the point), and then make lists of what the job applicant "did wrong." They then discuss and share their lists with the entire class.

3. Reinforcement. A mini-lecture by the teacher follows which emphasizes the key points of listening as demonstrated and discussed in the role-playing session above (e.g., eye contact, relaxed posture, and verbal following -- no topic jump). The microcounseling manual, Attending Behavior, is sometimes used to supplement the lecture.

4. Develop the model. Another role-play is held in which the job applicant does a more effective job (acts correctly). The other students observe and note the differences between the two sessions.

5. Practice. Students then practice this exercise in pairs within their groups so that the concept of attending behavior (listening) is experientially learned. Sometimes students practice in triads with the third

person acting as observer/evaluator.

### The Skills of Microcounseling

Numerous skills have been identified and field-tested within the microcounseling framework. These skills have been organized into a number of broad categories or clusters. First are the beginning skills of effective interpersonal communication: attending behavior, open questions, and minimal encouragement to talk. These skills help the learner to convey interest as well as to get the other person to talk and express him or herself more fully. They enable the learner to avoid disastrous early attempts at interpersonal communication. Although basic, these skills are essential for every person who wants to communicate with others. Indeed, experienced professionals, including teachers, also benefit from and welcome systematic microcounseling training in these basic skills.

Another cluster consists of selective listening skills, including reflection of feeling, paraphrasing, and summarization. These skills enable the learner to communicate understanding of affect as well as content of the other person's words. They also assure that both parties have the same understanding of what is being said. The need for these skills in numerous situations is obvious.

Some of the advanced skills include giving directions, expression of content, expression of feeling, self disclosure, interpretation, and direct mutual communication. Each of these microcounseling skills is described in a brief manual (Ivey and Authier 1978). These manuals are invaluable tools for the teacher who wants to describe microcounseling skills operationally to students.

The ESL teacher may not consider every microcounseling skill appropriate to the needs of ESL students. That is as it should be. Microcounseling is an open system. One ESL teacher might consider attending behavior as a critical need for all ESL students and incorporate attending behavior into the curriculum via microcounseling. Another teacher might determine to include selective listening skills as well as attending behavior. Each teacher could also elect to modify the existing microcounseling manuals to identify her/his own style. Each would have an individualized, yet systematic approach to teaching the concrete behavior

that she/he has identified as essential. The ESL teacher is encouraged to select those skills considered to be most important in a particular setting, or even develop definitions of new skills. Regardless of the specific skills considered most appropriate for effective functioning, the systematic technology of microcounseling may be utilized efficiently.

#### The Group "Micro-ESL" Format

Because ESL students often lack the lexical, syntactic, and socio-linguistic skills necessary to carry out adequate conversations in English, the microcounseling approach will be modified to include three preliminary language and culture components. There are, then, eight parts to a group "micro-ESL" exercise. The topic will again be "The Job Interview," and the interpersonal skill to be emphasized will again be attending behavior.

1. Vocabulary practice. The teacher analyzes the situation to be practiced (a job interview) and identifies the vocabulary and idioms necessary to carry out an adequate conversation. These are presented to the students.

Examples: application, interview, employer, employee, personnel, qualifications, hired, fired, laid off, resumé, references, position, opening, salary, over-time, union dues, a-month, an-hour, wage, to earn, to make, sick-leave, to bring home, after taxes, etc.

2. Grammatical structure practice. The teacher again analyzes the situation to be practiced and identifies the major grammatical structures necessary to comprehend and converse in a job interview situation. For example, an employer would commonly ask questions like:

"Have you ever been a cook before?"

"Where else have you been a cook?"

"How long have you been a cook?"

The teacher then knows that the students should have a good aural comprehension of yes/no and WH-questions in the present perfect. For other examples, the employee would probably respond with:

"Yes, I have." or, "No, I haven't."

"I've cooked in several restaurants," or,

"I was a cook at Golden Arches for two years."

"I've been a cook for two years," or,

"I was a cook in Cuba for two years."

Again, the teacher knows that it would be important for students to be able to orally produce the present perfect (affirmative and negative) and the simple past. Other structures the teacher may wish to review are: simple future (e.g., "Will I work weekends?" "You will make \$600 a-month.") Future progressive (e.g., "Will I be working weekends?" "You'll be bringing home about \$480 a-month.") The number of structures to be covered within one lesson would depend upon how advanced the students were, and on whether the grammatical structures being presented were new to them or a review.

3. Culture training. The teacher identifies the appropriate behavior in a job interview situation. It is important not to overwhelm the students with information, since they already have linguistic skills to worry about. Since it has been deemed important in our culture to demonstrate attending behavior during a job interview, the students are explicitly told how to show to an employer that they are listening. The teacher may find it useful to contrast any behavioral differences between the U. S. and the students' home culture. The teacher may also wish to use the students' native language here since the focus now is on cultural information and not language.

Example:

Appropriate Attending Behavior

- a. sitting with relaxed but attentive posture
- b. head facing interviewer
- c. looking at employer's eyes occasionally
- d. sticking to the topic, recognizing cues to respond

Inappropriate Attending Behavior

- slouching, bending over, or sitting too rigidly
- head down, away, or toward ceiling
- never/always looking in employer's eyes
- topic jumping, interrupting

4. Introduction. The teacher explains to the students that they are going to practice participating in a job interview, and that it is important for them to convey to the interviewer that they are listening. This introduction may be carried out in English or in the students' home language. The teacher may wish to quickly review some of the characteristics of appropriate attending behavior.

5. Training. The teacher selects volunteers to role-play an employment interview. The student playing the part of the applicant

is told to play his/her role as inappropriately as possible. The class is divided into small groups. After observing the role-playing session for 3-5 minutes, each group is to come up with a list of everything the job applicant "did wrong." The students should be encouraged to listen for content and previously-identified kinesic behavior, not grammatical perfection.

6. Reinforcement. The teacher presents a brief review of the key points of listening as demonstrated and discussed in the role-playing session above (e.g., eye contact, posture, and verbal following). Although use of English should be strongly encouraged, the teacher may wish to use the students' native language on occasion to make a point clear.
7. Developing the model. Another role-play is held in which the job applicant performs correctly. The other students observe and note the differences between the two sessions. Although this microcounseling exercise is intended to be carried out orally, if the teacher wanted to add a reading component, a Language-Experience Approach could be based on the dialogue used in the second role-play session.
8. Practice. Students are told to practice this exercise in pairs within their groups so that the concept of attending behavior is experientially learned. Students may practice in threesomes, with the third party acting as observer/evaluator. Perhaps the third person would be someone who lacked the linguistic skills or confidence to participate orally in the beginning.

#### CONCLUSION

Microcounseling is an approach to developing interpersonal communication skills which is used in many fields. By adapting the microcounseling format to include vocabulary, structure, and culture components, we too can use this approach to help our students communicate effectively. "Micro-ESL" is a systematic means by which to develop communicative competence. It is humanistic, does not require literacy, and it not only teaches language, but it also teaches about culture, basic survival, and vocational skills. The role-playing dialogues are interesting and will not lie outside the students' semantic skills. Students will participate equally and will experience success regardless of their degree of grammatical correctness. "Micro-ESL" can go hand-in-hand with any functional/notional syllabus, and it is an efficient and enjoyable way to educate adult ESL students.

REFERENCES

- Aldridge, E., and Ivey, A. 1975. The microcounseling paradigm in the instruction of junior high school students in attending behavior. *Canadian Counselor*, 9:138-144.
- Allen, D. 1967. Microteaching, in C. J. Cotrell and E. F. Hauck (eds.), *Educational Media in Vocational and Technical Education*. Columbus, Ohio: The Center for Research and Development in Vocational and Technical Education.
- Authier, J., and Gustafson, K. 1974. Using video to develop communication skills. *Biomedical Communications*, 2:10, 38:42.
- Bradley, C. 1977. Systematic interpersonal skill development for inner-city youth: A microcounseling approach. *Educational Technology*, 17(3):46-49.
- Chase, S., C. Doty, and C. Cotrell. 1971. Assessment of microteaching and video recording in vocational and technical teacher education: Phase IX -- Micro-Supervision. Columbus, Ohio: The Center for Vocational Education and Technical Education (ED 057 194).
- Chastain, Kenneth. 1971. *The Development of Modern-Language Skills: Theory to Practice*. Philadelphia, Pa.: The Center for Curriculum Development, Inc.
- Corder, S. P. 1973. *Introducing Applied Linguistics*. Harmondsworth: Penguin.
- Farid, Anne. 1976. Communication in the classroom: student-improvised dialogues, *TESOL Quarterly*, Vol. X, 299-304.
- Findley, Charles, and Nathan Lynn. 1980. Functional language objectives in a competency-based ESL curriculum, *TESOL Quarterly*, Vol. XIV, 221-231.
- Fox, James. 1978. TELEFUN: A pragmatic approach to functional learning materials development, *TESOL Quarterly*, Vol. XII, 297-309.
- Harrington, F. 1970. Development of a Self-Instructional Package on Cooperative Education Coordination Skills. Ann Arbor, Michigan: University Microfilms, No. 71-7474.
- Ivey, A. 1973. Media therapy: educational change planning for psychiatric patients. *Journal of Counseling Psychology*, 20:338-348.
- Ivey, A., C. Normington, W. Miller, and R. Haase. 1968. Microcounseling and attending behavior. An approach to pre-practicum counselor training. *Journal of Counseling Psychology*, 15: Part II (Monograph Separate) 1-12.
- Ivey, A., and J. Authier. 1978. *Microcounseling: Innovations in Interviewing, Counseling, Psychotherapy, and Psychoeducation* (2nd ed.). Springfield, Ill: C. C. Thomas.
- Johnson, Keith. 1977. The adoption of functional syllabuses. *The Canadian Modern Language Review*, 33, 5.
- Savignon, Sandra. 1972. *Communicative Competence*. Philadelphia, Pa.: The Center for Curriculum Development, Inc.

REFERENCES (continued)

- Sheraga, Mona. 1980. ESL with advanced high school students, *TESOL Quarterly*, Vol. XIV, 41-52.
- Stratton, Florence. 1977. Putting the communicative syllabus in its place, *TESOL Quarterly*, Vol. XI, 131-141.
- Taylor, Barry, and Nessa Wolfson. 1978. Breaking down the free conversation myth, *TESOL Quarterly*, Vol. XII, 31-37.
- Wilkins, D. A. 1976. *Notional Syllabuses*, London, England: Oxford University Press.

-----





ACTING METHODS APPLIED TO THE TEACHING  
OF ENGLISH AS A SECOND LANGUAGE

Stephen M. Smith

In this paper I will attempt to relate acting methods and techniques to aspects of learning a second language. I will take the position that actors and language learners share common goals (the most important being satisfactory performance) as well as certain obstacles to achieving them. The language teacher has a function analogous to that of a theatrical director--that of creating a setting and manipulating the actors in a manner which facilitates attainment of the central goal. I will then discuss various aspects of drama rehearsal which may be successfully incorporated into classroom teaching, in particular, ensemble work, the role of criticism, character study, role playing, and scene study.

One of the most important things that a language teacher must grasp early in his career is that the learner's emotional self and ego are sensitive things which must weigh heavily in his consideration of how to teach. Theatrical actors and directors also deal with ego and emotion to a great extent. Ego and self-esteem are on the line in the foreign language classroom as in the sometimes emotion-charged atmosphere of dramatic rehearsal. In both settings, the director and the language teacher are in a position to damage or preserve the exposed ego of the actors and language learners. The rehearsal is where the actor is called upon to sum up, then bare, a variety of real emotions. In this setting, he must cast away inhibitions which protect the fragile human ego, and be willing to go out on an emotional limb, take chances, be wrong, look silly, then try again. The director realizes that the actor's self-esteem must be preserved in order to facilitate continued honesty, openness, and flexibility in rehearsal, but at the same time, he knows that he must work with the actor to break down defenses which prevent him from being empathic, for empathy is a key to understanding the character to be portrayed as a real, living, feeling human being. Inhibitions are systematically reduced in rehearsal until all of the emotional valves are open and the honest emotions of the role are flowing. This is, of course, not an easy thing for the actor and director to achieve. Actors work for years to develop the ability to control inhibitions.

Language teachers can deal with the inhibitions of their students. The most obvious approach is to avoid making students' inhibitions any

worse than they already are. The language class, like the rehearsal, should be in an atmosphere conducive to open experimentation with the target language. The "director-actor" relationship must not be an adversary relationship. The language teacher should let the students know that he is on their side.

#### THE ENSEMBLE APPROACH TO TEACHING

The best plays are created by a "tight" ensemble. This means that the actors know each other well, and they trust each other. They seek advice from one another, and ask for feedback. The same goes for a group of language learners. They should feel free to criticize one another in a constructive manner, and they should learn to enjoy experimenting with the target language in front of their peers and teacher. Above all, a group of "actors" must not be afraid to be wrong in front of one another. "That's what rehearsals are for," the director will say. "Stick your neck out!"

Similarly, the teacher must foster the atmosphere of a tight ensemble. Language classmates must also be able to stick their necks out and drop their inhibitions to promote learning. The teacher should begin creating the proper atmosphere early in the semester. Class activities should emphasize students' names, personalities, and backgrounds. Coffee hours, beer seminars, pot-luck dinners, parties, and the like are more than extra opportunities to practice language in social settings. They can be crucial to the development of an ensemble work effort in the classroom. The teacher, like the director, must be ready to take time to shape a class into a good working ensemble.

In an ensemble approach to the class, the teacher is not only the director but also part of the ensemble. The teacher/director should promote an impulsive style. The impulsive language learner may make more frequent mistakes than the laboriously-reflective speaker, but he probably communicates more in the long run. The teacher can lead the class to use errors and mistakes as learning devices, but, as does the theatrical director, the teacher must read the students' reactions constantly, and be ready to turn the course of events to a more positive note, when things get too uncomfortable, so that the students never become too embarrassed and feel that they have failed in any sense. Proponents of the audio-lingual method went overboard on this point, by never allowing students to

attempt to use structures which hadn't been taught yet. They failed to recognize that the students' interlanguage is also a legitimate system of communication which they must be encouraged to use so that they can gain the confidence and experience necessary to develop a more efficient level of competence in the target language.

The student doesn't need to look good all the time, since this state of affairs doesn't exist in real life. The teacher can approximate life, by letting students get into trouble using the target language, and then by letting them see that they can get out of trouble through their own devices. Students, in role-play situations for example, may talk themselves into corners, unable to find the structures they're searching for, but they discover that by persisting and asserting themselves they can eventually make themselves understood. Role playing provides the students with safe opportunities to discover that even when they get in over their heads linguistically, there are ways out which don't result in total communication breakdowns. In this "safe" ensemble atmosphere, students develop confidence and learn to cope.

In role playing, "games," and scene work, the director/teacher should resist the temptation to correct too much and allow some mistakes to happen. Directors often take pains not to correct too much in drama rehearsals, realizing that role acquisition is a process which takes time. The actor needs time to work and rework, so as to discover, subsume, and prune elements of a character. But criticism and correction are important and need to be well placed for maximum efficiency. Exactly how to criticize depends on the teacher, the student, and the situation.

#### Criticism

Although it is not easy to specify what type of criticism is appropriate for a given situation, the role of criticism must be recognized. Language learners and actors are influenced by the consequences of their actions, and the rehearsals are where those consequences should be felt. For actors, rehearsals are working situations quite different from performance situations. Actors take safe, tried and true paths in most performances, but in rehearsals anything goes. The actor will experiment with a role, drawing criticism from the director. This is how discoveries are made. In rehearsal, the pressure to "perform" must be put at a distance. Language learners

commonly suffer from the opposite syndrome. They are too careful and reflective in the classroom, and steadfastly avoid exposing their ignorance and attracting criticism. Then, outside the classroom, they speak with reckless abandon, wantonly butchering what they have learned beyond recognition.

People, in the real world for the language learner, and the audience for the actor, seldom provide the same kind of feedback as the "rehearsal" can provide. Actors are no longer driven off the stage by angry tomato-throwing audiences when they give a bad performance. If an actor is bad, the audience suffers in silence. Similarly, in the real world, people don't always indicate to a foreigner when he is performing badly, either culturally or linguistically. To avoid a bad performance, actors must monitor themselves, and they count on monitoring from the director and fellow actors. From this regular monitoring in a controlled situation, they learn to monitor themselves more effectively. Frequently, a director waits until a scene or a rehearsal is over before verbally giving the actor the "notes" he has been taking. At other times, the director will frequently stop a scene in order to guide the actor through the reworking of it. These are referred to as "stop-and-go" rehearsals and require the utmost cooperation between actor and director, lest the actor become irritated with the director, who doesn't seem to be giving the actor enough freedom to accomplish what he is being asked to do. Directors sometimes go into great detail, and sometimes they don't. When they are silent, directors are usually waiting out what they see as a stage in the process of character development. If nothing develops, the director guides the actor through the stage, perhaps with a different approach. If the actor discovers something useful alone, the director tells the actor to "keep it. It works."

Monitoring can be encouraged in the ESL classroom in a number of ways. When, for example, in role plays or dialog-readings, attention is focused on intonation, students can be made aware of just how important intonation can be. A student may think he is expressing emotion X, when, because of his intonation, he is actually expressing emotion Y. The following exchange between a teacher and a student illustrates how, using a theatrical approach, the teacher can begin to instill self-monitoring in the student.

The student is midway through the reading of a dialog, when the teacher interrupts:

Teacher: Wait a minute! Excuse me, but why do you think your character just said to his girlfriend, 'I saw you with Tom last night'?

Student: Because she had told him that she couldn't go out on a date with him, because she was sick, and he wants her to know that she didn't get away with her lie.

Teacher: How does your character feel about her having lied to him?

Student: His feelings are hurt.

Teacher: Is he angry, too?

Student: Yes, of course.

Teacher: Ok. Good. I think you're right. Your character is hurt and angry. But when you said 'I saw you with Tom last night,' it sounded like you were greeting your milkman or waving to your paper-boy; you didn't sound like a hurt and angry boyfriend confronting his unfaithful girlfriend. Try it again, being conscious of intonation, and be careful of where you look and your facial expression.

Notice that in the above example, the teacher didn't tell the student how to achieve the desired results. There are many different executions of the line in question that would work. What the fictional teacher is attempting to do is to help the student "see" what it is he looks like, and to help the student realize that this type of exercise is more than a vocabulary reading exercise--it is an exercise in communication.

Giving the performer's classmates monitoring tasks actively involves them in the monitoring process, too. One approach would be to assign groups within the class to observe and report on different aspects of a role play dialog reading, or a video-taped performance, e.g., appropriateness of intonation, gestures, proximity, eye contact, vocal volume, pronunciation, word choice (if improvised), structure, and so forth.

It is important to allow time for the class to learn to work this way constructively and without inhibition. It must be approached as a collective intellectual activity, and not as a group attack on an individual's characterization and performance.

#### Character Study

Method acting, a theory developed by the great Russian director, Constantin Stanislavski (1936, 1949), is the technique of establishing the true emotions and motivations of a character, then producing those emotions

truthfully. The actor, like the language learner, is not as concerned with the words that come out of the mouth as with what those words mean to the speaker, why he chose those words, and what the words mean to all who hear them. The actor looks to the inner character for answers. To understand the inner character, the actor must seek to learn all the details of the character's life, his culture, background, philosophies, loves, hates and fears. Only after learning all these things about Hamlet can an actor truly know what is meant when he says, "To be or not to be."

A person learning to live in a foreign culture needs to become a better actor in much the same way as a stage actor does. Every good actor studies the characters he plays in depth. The foreign student in America should be motivated to study the American character. Character study for actors and cross-cultural participants should be a constant effort. Observation must be a habit if one is to become an expert at it. Chances are it will already be a habit for many who live in a foreign culture and who need to know what is going on around them. Students of cross-cultural communication could use some actor's techniques for focusing their energy so that their observations are productive. The "actor" wants to study the specifics of a character and should not be satisfied with generalities and stereotypes. Superficial, or "surface," manifestations of character should be the result of a thorough understanding of the "deep" character. It is not possible in limited classroom time to learn enough about the character of the target culture. Classroom time should be used to give actors and foreign students the tools with which to make every waking minute of their lives relevant to the task of character study.

The teacher can instill the capacity for observation in the students through guided classroom practice. It should be pointed out just how detailed observation can be, and how much information can be gathered through observation. To begin with, students should observe each other in the classroom. The teacher guides observation activities with questions. If, for example, the students observe that a classmate (chosen as a subject for observation) looks tired, the teacher can probe for specifics, e.g., Why does he look tired? How is he standing? How is he breathing? Are his eyes red? Is his coordination worse than it usually is? How late do you think he went to bed last night? To point out how much information is lost by not observing, students may be asked to face the wall and describe

what classmates are wearing. Even if they've been instructed to "look around the room" first, they will still have a difficult time describing details. The more they do describe, the more details the teacher demands. In this way, students learn how much information is missed in a task as large as "observation," when attention isn't focused. When the students are aware of what the teacher means when he says "observe," they are ready to make use of their skill.

Through the habit of observation, the student will be able to gather data to solve the daily problems of cross-cultural communication. Students and actors need to be aware of how others perceive their actions. They need to learn what it is they should do in order that those watching them accurately perceive the message they desire to communicate. Besides knowing what to do, they need to develop the flexibility to do it. They must be the master over their physical and vocal mannerisms. This is not to say that we want foreign students to attempt to become cloned Americans; the object is to give them the tools to express to people exactly what they mean and nothing more. This self-expression is tempered of course by the fact that we hope we are giving them background which motivates them to maintain an attitude of cultural-relativity. It often takes time to fully understand events witnessed in a foreign culture. We aren't training our students to express honestly on-the-spot every gut-emotion that comes to mind. We hope that through an understanding of the nature of people, including themselves, in cross-cultural situations, they will react to cross-cultural conflict with reserve and diplomacy until things become clear enough for them to make rational decisions about how they feel and how they should react to conflict, if any reaction is appropriate. In the meantime, they need the tools to analyze what is happening around them.

#### Developing Observation Skills

There are a number of "theater games" described in Spolin (1963) which are designed to develop observation skills that may be adopted for language teaching. They involve a variety of improvisational exercises and are designed to aid the actor in the creation of a particular role. The skills which these games are intended to develop center around communication and observation: observation of the self, others, and the environment, and how they interact. The following observation exercise for the ESL classroom is based on Spolin.

Where - Students are asked to name a place, for example, the living room. They are asked to recall everything they can about what a living room is, e.g., what is done there, and what are its physical characteristics? A list is kept by the instructor and everyone wracks their brains for details. For actors, the purpose is simple. How do they react when playing a scene in a living room, when there are no props or set to reinforce the living room image in the minds of the audience? The Second City Theatre performs almost exclusively with no set, yet in one night they take the audience all over the world and seem to leave very vivid impressions of where they are in any particular scene. For foreign students, such exercises strengthen their habit of looking at aspects of the environment that they might otherwise ignore.

Who - Students are instructed to observe people on a bus, or at a movie, or a shopping center, or anywhere they happen to be. They should select people who seem to be together and try to determine what their relationships are. If, for example, the student determines that two women in the front of the bus are mother and daughter, he should try to discover every single clue that gives the impression of a mother-daughter relationship. Later, the student communicates the minutest details of his observation to the class. The exercise can then be made more relevant to the teaching of culture by discussing cultural differences that might be at a play. For instance: are there any differences in the two women's appearances that contrast with the typical appearance of a mother-daughter team in the student's home country?

The key to observation exercises like the above is that they are deceptively simple. Detail is the name of the game. These are more than story-telling exercises. Players must break through the superficial. They must become better-than-average observers. Language learners, like actors, have a better-than-average reason to be talented observers.

Events are good subjects for observation. Students come to class with an account of something they observed. The verbal report should include all the details the student can recall and his analysis of them. The instructor then leads a discussion, seeking to clarify any confusing points about the incidents, and a role play could then be constructed, giving the students a chance to behave as the Americans did in the real-life situation. Motivations for the Americans' behavior should be established in the discussion. Students studying non-verbal behavior have an unlimited supply of subjects to observe. Eye contact patterns, gestures, physical proximity and other not-so-obvious manifestations of culture are keys to communicate within a new culture. Role playing and dialog reading can be used to bring any one of these paralinguistic elements into focus.



## Games

The types of "games" that are available for use in drama rehearsal and the ESL classroom are many and varied. They come from a variety of sources: group dynamics materials from psychology and education, the performing arts, our childhoods, and of course the instructor's/director's imagination.

Games are frequently used in drama rehearsal to tighten the ensemble. One of the main obstacles for actors and language learners to overcome is the "adult" reaction to expressing one's self truthfully. Feelings and emotions are suppressed. Dignity is at a premium in the adult world. Games, used wisely, can help adults rediscover the freedoms they knew as children. Games "allow" the adults to "play" in front of their peers. Through games, ensemble members become acquainted. The ensemble begins to operate as a unit. Its members begin to trust and cooperate.

Rehearsals often begin with a game or two in order to "warm up" the ensemble. The actors forget about the world outside, from which they have just come, and they concentrate on the group task. With the group energy focused, the work can begin. The following is a group warm-up which works well at the beginning of rehearsal or class.

Instructions: Members stand in a circle. Each thinks of a word. Going clockwise, each member says his or her word. Then the director chooses members at random to recall the word of some other specified ensemble member. The game is repeated with sentences. A flowing tempo is maintained.

The above game has multiple purposes: 1) it stops all activity which is irrelevant to the rehearsal/class; 2) it removes students from behind their desks, gets them out of passive sitting positions, and brings them fact-to-face with their classmates; 3) members learn each other's names; 4) everyone begins to listen in a very specific way, and to concentrate on the group; 5) the group begins to function as a unit; 6) good pronunciation becomes important; and 7) as the game becomes difficult, it becomes established that no one in the room is perfect--and that is a comfort to all.

"Role plays" are another type of "game." Role plays are not performances, but are for rehearsal and working situations. They give the actors the opportunity to explore alternative solutions to characterization

problems without the pressure of entertaining an audience. One goal is to expand the actors' vocabularies of behavior, and increase their comprehension of a wider range of possible behavior in others. Role playing, accompanied by feedback from the director/teacher and other ensemble members, helps actors to learn to distinguish between what they feel is, or should be, happening, and what really is happening. Actors learn to know themselves, and they try to develop an awareness of how others perceive them. Role plays can be designed or adapted to isolate and illustrate a variety of teaching points, whether it is to teach culture, give students practice with new intonation patterns, practice conventional speech formulas, emphasize non-verbal behavior, etc. The actors have an opportunity to "ad lib," reinforcing interpretation and role-understanding.

Ensemble members can contribute by correcting the actors' behavior in a role play exercise. If the actors' words are inappropriate, they can be "corrected," and the scene replayed until the actors have a feel for what is appropriate. A dialog reading becomes a role play when the actors are directed to lay down their scripts and improvise the unmemorized scene. The actors then use their own words to express the thoughts in the scene.

The following example of a role play designed for the ESL class is very controlled, and since the students are given specific tasks, it is not difficult. It provides experience interacting appropriately in a gift-giving situation. The students learn what is "right," and how it feels to be "right." Hopefully, they will develop a feel for why the American way of gift-giving is fun and expressive. Because they have learned how to express themselves to Americans when in a gift-giving situation, our students will be able to enjoy themselves, as Americans do, and it is less likely that they will be bewildered as to why the Americans don't just behave as the foreigner's countrymen do.

### Role-Play

This exercise might be useful in an ESL classroom briefly before the Christmas gift-giving season approaches, so that foreign students have an opportunity to learn something about the American way of giving and receiving gifts. The students benefit from learning how the American customs differ or resemble their own customs, as well as the opportunity to "rehearse" the roles they may be expected to "play" during the Christmas season.

Objective: To aid foreign students in learning to participate

comfortably in the American ritual of gift-giving.

Situation for role play: Small group of participants is gathered at the home of an American and presents him with gifts. Perhaps it is his birthday or a going-away party. The American opens the gifts one at a time, reading each card first, and reading the funny cards out loud, or in some way displaying enthusiasm for the message on the cards. The recipient opens each gift immediately after the card which accompanies the gift. Enthusiasm is expressed for each gift, although they are obviously different in value. Each guest at the party observes the opening of the gifts and says something brief as his or her gift is opened regarding why she/he thought that particular gift was appropriate for this occasion: (I thought of you the minute I saw this, or, I knew you would need a good book to read during your long flight, etc.). After the gifts are opened, the exercise is ended, and the actors as well as the observers discuss the actions of the cast, and why the situation was different than the way things would be in the native country of the non-Americans present.

This could be played again, giving a different actor the chance to play the recipient of the gifts. There are a variety of interesting alternatives: 1. This could be carried out first with a real American as the recipient, and then after the discussion, the game could be replayed with the foreign participants imitating the American's recipient-behavior. 2. A foreign student could be the recipient the first time, with instructions to guide his behavior. Then, after discussion, the American could play the recipient, and the group could discuss the differences in recipient behavior.

### Scene Study

Theatre production courses in which students rehearse short plays and then perform them, have been implemented successfully in the past. Via (1976), for example, gives an informative account of his experiences with theatre classes for ESL students at the University of Hawaii's East-West Center, and it is apparent that his approach could be adapted to suit a variety of ESL purposes both in the U. S. and abroad.

Short one-act plays, or a variety of scenes from different plays could be produced in one semester. Naturally, the best scripts are those with natural spoken language that is not dated or highly stylized. The director could implement "ensemble-tightening" experiences, allowing the students time to read a variety of roles from a variety of scripts, and to take part in role-play improvisations.

Rehearsals begin with group vocal "warm-ups" of a musical nature which gets the students' physical apparatuses flexed and moving, and gets the group energy focused as the students begin to listen to one another in order to perform the warm-ups in unison. Light, relaxation-

oriented physical warm-ups are taken quite seriously by actors. They serve to relax the actors mentally as well as physically. They also free the actors from the defensive fortresses of desks, chairs, and passive sitting positions.

Work on the script begins with a "read-through." The group then discusses the script in detail, seeking to psychoanalyze the characters. "Blocking" comes next. The director gives the cast directions as to when to enter, where from, and where to stand. The director can let the actors be creative with much of their own movement, but he should coordinate blocking to keep it organized, and to give the actors motivations and a few acting crutches when necessary. When actors are still in the process of discovering what a character is all about, they sometimes appreciate being given something to do with their bodies. The director can help by suggesting how to use the set and props to make the scene real. Many ideas occur to a person watching a scene which the actors don't have time to think about.

Throughout the rehearsal period, constant questioning of the script by everyone, especially the director, is crucial. What does such-and-such a line mean? Who is being referred to in conversation X? Why did character X say that about character Y, and what does that tell us about the relationship between X and Y? In short, a lot of time should be devoted to reading between the lines.

Line memorization should not be pushed by the director. Actors will find it easy toward the end of the semester after working on the same piece repeatedly. By then, the script will be solid meaningful thoughts in the actors' heads. Thoughts are easier to remember than words.

Finally, at the end of the semester, the scenes should be presented to an audience. The performance is the final step that makes all the concentrated work on the limited sample of the English language meaningful. In addition, a dimension is now added to the notion of communication, because now, the actors must not only communicate with each other, they must communicate to the audience. All of the decisions that have been made throughout the semester about how to interpret the script now must be seen clearly by an audience that will watch the action only once.

## Conclusion

In this paper, I have attempted to show that theatre methods can be injected into the ESL classroom in varying doses: either as part of a teacher's general philosophy, or as supplementary materials, or as a complete semester-long theatre production course. It is not my purpose to prescribe a neatly packaged curriculum combining theatre and language-teaching methodology. But in the application of an enlightened eclecticism to the choosing of ESL teaching strategies, teachers have a great deal to gain by borrowing selectively from one of mankind's oldest means of expression: the performing arts.

-----

REFERENCES

- Spolin, Viola. 1963. *Improvisation for the Theatre*. N. Western University Press.
- Stanislavsky, Constantin. 1936. *An Actor Prepares*, by Elizabeth R. Hapgood, New York Theatre Arts Books, New York, N. Y.
- Stanislavsky, Constantin. 1949. *Building a Character*, by Elizabeth R. Hapgood, New York Theatre Arts Books, New York, N. Y.
- Via, Richard. 1976. *English in Three Acts*. University of Hawaii Press.
-

REVIEWS

READING IN A SECOND LANGUAGE: HYPOTHESES, ORGANIZATION AND PRACTICE.  
Ronald Mackay, Bruce Barkman, and R. R. Jordan, editors. Rowley, Mass.:  
Newbury House Publishers, Inc., 1979. Pp. v + 208

Mackay, Barkman and Jordan have collected a series of readings, most of which have appeared as journal articles, in Reading in a Second Language: Hypotheses, Organization and Practice. The collection is divided into three sections. Section One, "Hypotheses," forms the book's theoretical base. The editors note that the following authors accept Kenneth Goodman's (in Singer and Ruddell, 1976) "psycholinguistic guessing game" premise, that "the reader takes advantage of the redundant features of language to reconstruct an author's message from a text" (p. 1):

Chapter 1: James Coady: "A Psycholinguistic Model of the ESL Reader." Coady summarizes Goodman, and argues in favor of a processing strategy shift (from grapheme-phoneme to lexical and ultimately contextual meaning strategies) as an ESL reader becomes more proficient.

Chapter 2: Colin Harrison and Terry Dolan: "Reading Comprehension--A Psychological Viewpoint." The authors performed two factor analytic studies to seek an underlying taxonomy of reading skills. They located, rather, a single factor, prompting them to take a unified group-oriented "vital response" approach to the teaching of reading.

Chapter 3: Muriel Saville-Troike: "Reading and the Audio-Lingual Method." Saville-Troike calls for recognition of the separateness of reading skills, and offers a series of suggestions for doing so (e.g., recognizing sociocultural aspects of reading, and using morphemic and contextual clues).

Chapter 4: Donald Sim and Marsha Bensoussan: "Control of Contextualized Function and Content Words as it affects EFL Reading Comprehension Test Scores." The authors conducted a study to determine the relative influence of content vs. function words in reading test comprehension. No significant difference was discovered, leading the authors to conclude that function words (which yield understanding through cohesiveness) are at least as important and "need to be taught and tested to the same extent as content words." (p. 40)

In Section Two, Organization," the authors offer readings to bridge the theory/teaching gap. Based on Goodman's call for acknowledgement of

reading complexity, the vast matrix of readers' needs and skills are incorporated into several suggested organizational schemes.

Chapter 5: Mark A. Clarke and Sandra Silberstein: "Toward a Realization of Psycholinguistic Principles in the ESL Classroom." The authors, who helped write Reader's Choice (Ann Arbor, University of Michigan Press, 1977) assert that fluent reading is purposive reading which draws on a multitude of factors, e.g., the learning environment, the role of the teacher, materials evaluation, preparation and usage, and language skills. This chapter is an excellent nutshell of the beliefs on which Reader's Choice was founded.

Chapter 6: David E. Eskey: "A Model Program for Teaching Advanced Reading to Students of English as a Foreign Language." The author outlines a three-level diagram of reading as it relates to language (meaning, form, The Printed Page), and an intensive (focused)/extensive (general) reading model based on that diagram. The resulting fusion is quite appealing, especially his discussion of contrastive rhetoric.

Chapter 7: Ronald Mackay: "Teaching the Information-Gathering Skills." This article is an appeal to recognize the vast complex of linguistic markers which signal the transmission of information. For example, Mackay presents a table of discourse markers and the informational notions they express (e.g., Additive/Similarity: "equally, likewise, similarity," etc.). He concludes with suggestions for inclusion of such signals in reading instruction.

Chapter 8: Sheila Been: "Reading in the Foreign Language Program." The writer here offers an analysis of reading theory, covering its aims, classroom practices, and analytic procedures. She closes with a strong argument in favor of teaching reading by separating "reading for language" (reading aloud, vocabulary work, and literal comprehension exercises) from "reading for meaning" (context support and cues to aid the reader in ignoring the linearity of reading).

In the third and final section, "Practice," the editors present applications of modern reading theory to the classroom:

Chapter 9: Ronald Mackay and Alan Mountford: "Reading for Information." Focusing on the needs of students of English for Special Purposes (e.g., science and technology), the authors closely analyze the skills necessary and the problems often present in reading for special information. Their analysis and suggestions are both linguistically specific (e.g., cohesive



reference: "this case, those points," etc.), and broadly analytic (e.g., knowledge of rhetorical structure) with many illustrative examples.

Chapter 10: John Munby: "Teaching Intensive Reading Skills." After a crucial distinction between skills (e.g., reading for required information, reading for implied meaning) and nonskills (e.g., character and plot) study, the author presents a coherent and broad-based intensive reading program to train skills study. He uses a long example from a novel which could lend itself to nonskills study as well: Chinua Achebe's Things Fall Apart. His multiple-choice format stresses analysis of distractors.

Chapter 11: Mary Eleanor Pierce: "Teaching the Use of Formal Redundancy in Reading for Main Ideas." The author here notes that "advanced reading, particularly at the college level, requires the ability to recognize and relate a series of ideas." (p. 159). She proposes that such recognition involve formal rhetorical cues, of which she gives many examples (e.g., the relation of facts to conclusions, and the nature of the paragraph "environment" [p. 163]; i.e., the topic sentence and its relation to other sentences in the paragraph).

Chapter 12: Ruth Berman: "Analytic Syntax: A Technique for Advanced Level Reading." Building on the notion of "structural paraphrase ("reworded and juggled" phrases with a minimum of content change [p. 180], the author presents a pedagogical scheme of types of paraphrase for students to use, for example, paraphrase of pronominal reference, sentence connectors, and negation. The approach is analytic; it aims at getting students to dissect and study the syntax of what they read.

Chapter 13: Salwa Ibrahim: "Advanced Reading: Teaching Patterns of Writing in the Social Sciences." The author, similar to the argument of Pierce, above, stresses overall flow of idea. He presents a teaching scheme, with a sample passage, in which the support structure of an article on Herbert Spencer is analyzed using diagrams and flowcharts.

There are two objections which can be raised about this book. It demonstrates a lack of understanding of reading theory advancements after Kenneth Goodman, and it lacks the integration necessary to be of practical value to the classroom.

The theoretical base for almost every article in this collection relies heavily on Smith (1971) (1973), or Goodman (op. cit.). These two authorities,

especially Goodman, have been shown to lack explanatory power, cf., Gibson and Levin (1975). Recent integrative theories, for example, those of Rummelhart (1972)(1978), and Thorndyke (1977), have challenged Goodman and Smith's top-down, expectation-oriented views. Kintsch and Van Dyk (1975) and Kieras (1978) have offered empirical support for such integrative theories. More generally, reading research is, and should be, in a state of investigative flux. Yet there is little mention in this volume of pressing research concerns, such as how the organization of a text affects reading comprehension, for which Lackstrom (1974) has proposed a solution. Indeed, the articles in this collection make little mention of direction that reading research should take. Rather, the Smith/Goodman stance is taken as a "de facto" explanation of what happens when an individual learns to read in a second language.

In addition, when one considers the most likely user of this book, the classroom practitioner, there is a lack of integration of the articles presented. Overlooking the abovementioned theoretical difficulty, when any collection such as this is placed on the market, it should include an integrative challenge to the reader. This would provide a skeletal framework from which to operate in the classroom. Such a framework could readily be accomplished by the inclusion of integrative questions after each article. Within the Goodman/Smith theoretical framework of this volume, such questions would have been possible. For example, after Chapter 13: What is the relation of the stance expressed by Ibrahim to that of Pierce? In the classroom, how might this stance need to acknowledge cross-cultural contrastive rhetoric research, which Eskey notes? Another alternative would be a fourteenth chapter, written by the editors, noting how the articles interrelate.

To conclude, if the theoretical bases of these articles had been updated, and the articles integrated to challenge the classroom user, the potential for this collection as a workable and satisfying pedagogical whole would have been greater.

-- Fred G. Davidson

-----

REFERENCES

- Gibson, E. J., and H. Levin. 1978. *The Psychology of Reading*. Cambridge, Mass.: MIT Press.
- Lackstrom, J. 1974. The comprehension of English for science and technology arguments and definitions. In J R. Cowan (ed.), *Language for Special Purposes* 2:1, 45-66.
- Kintsch, W., and T. A. Van Dyk. 1975. Comment on se rapella et un résumé des histoires. *Languages* 40, 98-116.
- Kieras, D. E. 1978. Good and bad structure in simple paragraphs: effects on apparent theme, recording time, and recall. *Journal of Verbal Learning and Verbal Behavior* 17, 13-28.
- Rummelhart, D. E. 1972. Understanding and summarizing brief stories. In D. Le Berge and J. Sammuels (eds.) *Basic Processes in Reading and Comprehension*. Hillsdale, N. J., Erlbaum.
- Rummelhart, D. E. 1978. Notes on a schema for stories. In D. G. Bahrow and A. Collins (eds.). *Representation and Understanding: Studies in Cognitive Science*. New York: Academic Press.
- Singer, Harry, and Robert Rudell. 1976 (eds.) *Theoretical Models and Processes of Reading*, 2nd ed. Newark, Delaware, The International Reading Association.
- Smith, Frank. 1971. *Understanding Reading: A Psycholinguistic Analysis of Reading and Learning to Read*. New York: Holt, Rinehart and Winston.
- Smith, Frank. 1973. *Psycholinguistics and Reading*. New York: Holt, Rinehart and Winston.
- Thorndyke, P. 1977. Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology* 9, 77-110.

-----



IMPROVING SPOKEN ENGLISH. Joan Morley. Ann Arbor, Michigan: The University of Michigan Press, 1979. Pp. xviii + 349.

Improving Spoken English is an ESL pronunciation textbook designed for high-beginning through intermediate level students. This book is the first of a two-part text. There are two units in this volume: Unit One deals with suprasegmentals (stress, rhythm and intonation); Unit Two, which deals with vowels, gives the student articulatory practice and exposes him to spelling information. After these two units, there is a section of review and practice material. Finally, the Answer Key/Teacher Script provides a transcript of the tapes which go along with the text, so that the teacher may read the oral practice material if that is more convenient, or if a language lab is not available.

Morley's text has a number of outstanding features: First, its concern for developing skills which will carry beyond the classroom; second, its emphasis on vowel reduction in phrases; third, its treatment of vowel sounds; and fourth, its supplemental material and general sensitivity to phonological variability.

One strength of Morley's text which makes it unique is that it encourages independence on the part of the students. For example, active use of the dictionary, a skill which transfers to the world outside the pronunciation classroom, is promoted through the assignment, in each vowel lesson, of six words which are to be looked up and their pronunciation copied exactly. Also, Morley makes some attempt to deal with the prediction of pronunciation from spelling, which ideally should be one goal of pronunciation teaching. Finally, Improving Spoken English emphasizes self-monitoring of pronunciation. An awareness of the physical aspects of pronunciation is stressed. For example, students are expected to own small pocket mirrors for use in pronunciation class. Schematic drawings showing vowel height and tongue position are provided for each vowel. Students are instructed to watch the teacher, try to reproduce the mouth and jaw movements, and note how it feels to pronounce each vowel.

Another strength of the text is the attention it rightly gives to stress timing, that is, the rhythm of alternating accented and unaccented syllables and the concurrent reduction of unstressed vowels. The text would be stronger, however, if the student were taught more explicitly where to place the main stress of a phrase or utterance.

The treatment of English vowels is good. All of the necessary contrasts are presented in Unit Two. For each lesson, Morley provides a key word and a vowel number. Phonetic symbols (slightly adapted from the International Phonetic Alphabet) are presented but not emphasized. Instead, Morley prefers vowel numbers. In her opinion, the use of vowel numbers helps remind the student not to merely use allophones from his native language in his English pronunciation. A problem with the use of vowel numbers, however, is that it does not help the student relate spelling to sound. Morley does present spelling-pronunciation correspondence patterns individually for each vowel (e.g., /iy/ is predicted by eC<sub>2</sub> [p. 132]; /ey/ is predicted by aC<sub>2</sub> [p. 136]; /uw/ is predicted by uC<sub>2</sub> [p. 163]; /ow/ is predicted by oC<sub>2</sub> [p. 168]; and /ay/ is predicted by iC<sub>2</sub> [p. 212]), but it would be more helpful if her vowel symbols lent themselves to general systematized spelling patterns.

Finally, the practice materials in Unit One, as well as those in Supplements B and C, are extensive and useful. They are presented against the background of the suprasegmental material from Unit One. Throughout the vowel material, Morley has taken into account phonological variability. For example, in an /æ/ - /a/ contrast drill (pp. 283-284), there are no a's before voiceless fricatives or before nC, where some educated speakers vary between /æ/ and /a/ for the same word. In a /uw/ - /u/ contrast drill (pp. 289-290), there are no oo's before labials, where some speakers vary between /uw/ and /u/ for the same word. Improving Spoken English does present the /a/ - /ɔ/ contrast, but directs the teacher to make his own decision about whether to try to teach it (footnote, p. 172). In other words, phonological variability in this area is acknowledged, and it is left up to the teacher to decide how to present it to the students.

There are certain deficiencies in the text which should be noted because they require the teacher to go to other sources for supplementary material. Furthermore, certain matters of presentation are problematic.

The most obvious gap in Morley's text is that consonants are not covered. This is because Improving Spoken English is only the first part of a longer text. Intermediate Spoken English, the second part now in preparation, will cover consonants. Other major gaps are in the area of suprasegmentals. Unit One introduces intonation. But, while the coverage of utterance-final intonation is fairly thorough, utterance-medial intonation

is missing, except for the treatment of phrases in a series. A serious problem with Morley's treatment of utterance-final intonation is that the student is not specifically taught where to place the peak in the intonation pattern. The peak depends on the placement of major utterance stress, and, as mentioned, the topic of utterance stress is not covered.

Several matters of treatment and presentation create difficulties, too. Unit One briefly introduces morphology. In this section, the student is given a heavy load all at once; he is presented with the voiced/voiceless consonant distinction at the same time that he is expected to make use of it in regular past tense and plural formations.

In ESL pronunciation teaching, a recurring problem, in terms of phonological variability, is postvocalic /r/. Morley's coverage of phonological variability for vowels preceding /r/ is quite complete except for the /æ/ variant in words like Mary and carry, which she does not acknowledge. Her presentation of postvocalic /r/ is interesting, but seems a bit contrived. /ɜ̃/ is referred to in three ways, as a high central vowel (p. 187), a syllabic r (p. 190), and a vowel-r (p. 189). We are told that the vowel of the vowel-r is silent (p. 187). The word, bird, for example, is actually pronounced 'b'rd.' Morley claims that /ɜ̃/ is the sound which follows all other vowels (except /æ/) when r occurs postvocally. Obviously, /ɜ̃/ is not supposed to add another syllable, but that is the implication. Though there may be no way to deal simply and in a straightforward way with postvocalic /r/, Morley's treatment of it is contradictory and confusing.

In summary, Improving Spoken English is a good ESL pronunciation resource, particularly in its effort to develop a resourceful and self-critical student, and in its presentation of vowel sounds. But the text must be used in combination with other teaching materials in the areas of consonants, phrase and utterance stress, and intonation.

-- Mary Siekert

-----







# *TESL STUDIES, 1981*

- LAWRENCE F. BOUTON *The Imperative Tag: What Is It Really Like?*
- CAROL CHAPELLE  
JOAN JAMIESON *ESL Spelling Errors*
- LONNA J. DICKERSON *Evaluating, Selecting and Adapting Pronunciation Textbooks: Guidelines for ESL/EFL Teachers. I. The Evaluation and Selection Process*
- WAYNE B. DICKERSON *A Pedagogical Interpretation of Generative Phonology: II. The Main Word Stress Rules of English*
- STEPHEN B. DUNBAR *Models of Human Memory in Second Language Learning*
- JOAN E. FRIEDENBERG  
CURTIS H. BRADLEY *Micro-ESL: A Method for Teaching the Functional/Notional Syllabus and Developing Communicative Competence*
- STEPHEN M. SMITH *Acting Methods Applied to the Teaching of English as a Second Language*

## REVIEWS

- FRED G. DAVIDSON *RONALD MACKAY, BRUCE BARKMAN AND R.R. JORDAN, editors: Reading in a Second Language: Hypotheses, Organization and Practice*
- MARY SIEKERT *JOAN MORLEY: Improving Spoken English*











UNIVERSITY OF ILLINOIS-URBANA



3 0112 038143233