A Preliminary Study on the Development
of the Test of Oral Proficiency

SHIRO SATO

Introduction

Surrounded by an atmosphere of ever-increasing opportunities to hold international-level meetings and conferences in and out of Japan, more people are beginning to point out the urgent need for Japanese to express, in both spoken and written form, English much more clearly and effectively than before. I happened to find a rather shocking statement in a book which says : Japan as a nation has been and will be losing astronomical sums of money due to each delegate's almost total inadequacy to express himself in English, particularly in behind-the-curtain talks with the delegates of other nations. The factors leading to such inadequacy may consist of deep-rooted non-linguistic knowledge (i.e. logical thinking, paralinguistic knowledge, etc.) as well as linguistic competence. Though there may be a case where lack of non-linguistic knowledge plays a critical role in communication, it is perhaps linguistic competence in most cases that plays a decisive role. Thus, classroom teachers and school administrators should make the utmost effort to have their students become equipped with linguistic competence ; ignoring this may lead to students' acquisition of surface level knowledge of the target language. With this in mind we shall first examine whether the constant use of the traditional teacher-made oral examinations of various types will ultimately lead to the students' acquisition of linguistic

competence in speaking skills. After this examination we shall further elucidate several indispensable factors to be employed in developing oral proficiency tests in the future. The author holds the view that those factors should play crucial roles in determining whether a given oral proficiency test is valid. In view of the fact that research in this field is still in the pioneering stage I should like to limit the scope of the present study to "Validity" and therefore leave the other two essential factors "Reliability" and "Economy" to future study.

The Validity of Achievements Tests to be Used as a Basis for Developing a Test of Oral Proficiency

It is assumed that most of the oral tests conducted in a classroom follow a "discrete-point approach," and so long as these tests attempt to uncover the degree to which the student has learned within a semester, they may be valid and worth using. But the appropriateness of the "discrete-point approach" to be used in developing an oral proficiency test has been challenged seriously both on theoretical and empirical grounds.

The "discrete-point approach" assumes that mastery of the sum total of specific linguistic knowledge or skills will ultimately lead one to acquiring overall proficiency of the language, and that such an approach enables the language teacher to list all the linguistic items in each skill and test them. A major framework of this approach was shaped by Lado (1957), though as early as 1945 Fries gave the impetus to Lado's subsequent theorizing. The two advocates and their followers stress the importance of comparing the target language with the native language. The comparing of the two languages are primarily done to determine the areas of difficulty which the

second language learners might encounter in learning a second language. A point by point analysis of each component of phonology, morphology, and syntax of the two languages is made so that the investigator can analyze and predict the problem areas. Those areas where the target language differs most radically from the native language will be the most difficult and conversely, those areas where the two languages are found similar will be comparatively easy for the second language learner. The hypothesis underlying this theory is that the second language learner tends to transfer the forms, meanings, and the distributions of forms and meanings of his native language into those of the target language when attempting to produce and comprehend both the spoken and written messages. The acquired habits of the native language are so deeply set in his nervous system and his muscular, intellectual and emotional processes that any minute shift of habit from native to second language becomes extremely difficult. Consequently, it becomes evident that the greater the discrepancies of the structure in question between the two languages, the more difficult it will be for the learners and the smaller, the less difficult.

The assumption described above is, in a similar way, employed in testing ; to put it in simpler terms, our task is to compare the target language with the native language and find the trouble spots, and test them.

We shall now examine the validity of this approach if it can be used as a basis for preparing test items of the proficiency type test. The first and foremost contention against this approach would be that it almost totally excludes the possibility of trouble areas inherently existent in the target language itself. It is quite conceivable that the learner not only transfers the habits attributable to his native

17

language, but also, in much the same fashion, does transfer the habits traceable purely to the target language itself. The hypothesis of this suggestion is that there should be areas where the comparison on all levels of structure (in a broader meaning) between the two languages is impossible; therefore, when the second language learner faces a structure which does not exist in the native language, it is obvious that he is forced to use the strategy other than the transferring of the rules he has learned in the acquisition of his native language. Furthermore, it is suggested that the more proficient the learner becomes, the less he will have to resort to the rules acquired in his native language; rather he will analogize or generalize the rules established in the course of second language learning.

Though overlapping with the first contention, the second question posed against contrastive analysis is that there leaves a considerable amount of doubt as to the possibility of comparing all components of the two language systems; it seems to me that, provided it is possible, the comparing of phonology might be the easiest and thus may be possible, syntax the next easiest but lexicon the most difficult of all (maybe totally impossible). In fact, no full description of this has yet been made available.

Added to the difficulty of comparing three levels of language systems is the task of stating the criterion by which the degree of difficulty is clearly defined. On what basis can we identify that the two structures compared have the identical forms, meanings or appear in the same distribution.

Finally, even if the appropriate description which defines the level of difficulty is made, such a description has to undergo constant changes as the learner makes progress in second language learning. Thus, it must be admitted that the contrastive analysis hypothesis is

invalid.

Among many experiments which attempt to prove that the contrastive analysis hypothesis alone does not fully explain second language acquisition, only a few representative ones will be discussed. Most of these experiments are designed to analyze the types of errors that the subjects actually make and categorize in types. Such error types are typically categorized into two : interlingual and intralingual (the former refers to errors ascribable to the native language and the latter to the second language itself). Another concern of these investigators is to find out the relationships between the level of the learners (i.e. the beginners, intermediate, and advanced) and the strategies each level of students seek.

Taylor (1975) conducted an experiment with twenty Spanish students of English as a second language, asking them to write eighty English sentences after hearing an equivalent number of Spanish sentences. The results showed that although both interlingual and intralingual errors made by the elementary and intermediate subjects did not differ qualitatively, they differed quantitatively. In other words, both levels of learners used interlingual as well as intralingual strategies but that the primary level subjects relied more heavily on interlingual and the intermediate level subjects on intralingual. The data indicated that as the proficiency of the learner increased, the reliance on intralingual stategies also increased proportionately. A similar finding resulted from the research by Dommergues and Lane (1976) who analyzed the errors made by 438 French freshmen at the University de Paris in the syntax subtest (consisting of 40 items) of the TOEFL. They report the interesting data :

The more students acquired a global knowledge of English,

measured by their percentile rank on the five subtests of TOEFL combined, the less they were trapped by the items that predominantly tapped interference.[1]

It is conceivable that when any one attempts to learn new knowledge, he tends to depend on knowledge he has already acquired. It follows, then, that for the very beginning stage learners, the type of knowledge on which they can rely is almost exclusively that of his native language, but the intermediate stage learners who have already acquired a considerable amount of English have a broader range of stock whereby they can utilize the knowledge of the target language itself in generating acceptable English utterances. It is thus
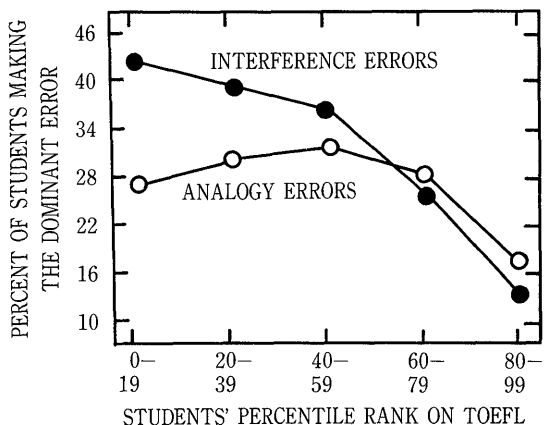

FIGURE I. The Percent of 438 French Students Who Erroneously Accepted an Ungrammatical Sentence on the TOEFL Syntax Test as a Function of the Students' Global Mastery of English. Eight Items that Were Independently Predicted to Elicit Errors of Interference and Seven Predicted to Elicit Errors of Analogy Are Plotted Separately. (Here "interference" refers to interlingual and "analogy" to intralingual.)

quite possible that these learners apply their acquired rules of English first and when they feel those rules will not work, they apply the rules of their native language If we extend this assumption a little further, it appears foreseeable that the advanced students rarely commit the interlingual type errors although generally they make considerably fewer errors. This assumption was correct when Dommergues and Lane (1976) presented the following figure.[2] The foregoing experimental results and discussion on the tendency of errors made by the different levels of learners are aptly illustrated in the same figure.

An easily recognizable fact on glancing at this figure is that even the primary stage learners use intralingual strategies to a great extent though to a lesser extent than their counterpart learers. This may suggest that upon learning a second language the second language learners start to analogize, systematize, and regularize the input data of the target language to which they are exposed. Also, the data shown in the above figure seem to indicate that learning a second or foreign language involves a creative process on the part of the foreign language learners in understanding and expressing the intended message. Persuasive data on this were disclosed in Lapkin and Swain's study (1977) with bilingual (English-French and French-English) and unilingual (English and French) subjects. They reached the conclusion that as a result of error analysis of cloze tests, both groups of subjects, bilingual and monolingual, there were no significant qualitative and quantitative differences found in recurrent errors between the two groups indicating also that interlingual errors can not be the only cause of errors.

The foregoing discussion leads us to admit that we have no assurance that the student who has consistently shown high scores on classroom tests (achievement tests) will also demonstrate high

21

scores on proficiency tests. More precisely, the student may have acquired specific knowledge or linguistic components as demonstrated in an achievement test but may not have yet reached the stage where the full control of other linguistic components as well as nonlinguistic competence are called for. In short, while the results of achievement tests based on "discrete-point approach" will give us information of how much short-term objectives have been met, they may give us little or no indication of how much long-term objectives have been met. If the teacher has continued using achievement tests throughout the course and later found that his students have attained little linguistic competence, it would be far too late to suddenly reshape the students' past linguistic experience into a desired language behavior.

For the reasons stated above it is necessary for the classroom teacher to use a proficiency type test to periodically check the student's linguistic competence in a functional context and, by doing so, the whole educational process would become more meaningful.

It seems clear that a person cannot gain proficiency without having acquired a certain amount of linguistic knowledge during the course. It is this knowledge that an achievement test can and should measure. Certainly a "discrete-point approach" will be an useful and effective tool to be used in making the achievement test successful. But since our major concern is with how successfully each student has incorporated his knowledge of grammar, vocabulary and pronunciation in his speaking skills we shall have to seek for a totally different framework from which the test items are constructed. And such a framework must be the one which attempts to uncover the testee's communicative ability in speaking skills.

Discussion

Several techniques have been employed, so far, in trying to tap one's communicative competence : translation method, spontaneous speech, oral reading, directed speaking tests, interviews, free conversation, etc. Before we adopt any one of these techniques we should always examine what a test purports to measure and check to see if the test gives us information we seek. Most communication tasks require us to communicate between or among persons and each communcation takes place in a particular socio-cultural situation. It follows, therefore, that the type of information we need will be the one which tells us how well the testee can perform orally in a particular socio-cultural context. The careful examination of each technique cited above against this basic criterion will automatically exclude several techniques at least theoretically : indirect speaking tests, oral reading, translation method, spontaneous speech, directed speaking tests. Directed speaking tests may become valid if the directions are framed in such a manner that will enable the testee to respond orally only after he has grasped the socio-cultural situational meaning. The problem of this technique is that because of the nature of the test, the testee might be able to obtain a high score whereas he might perfectly fail in communicating if the other party joins in a conversation which incorporates a social context. This will leave us interviews and free conversation. At a first glance both techniques seem effective in tapping one's competence in speaking skills. It must be pointed out, however, that prior to the testing, a careful preparation is needed ; without this, both techniques may, at best, end up showing the list of scores based on the evaluators' subjective judgment. Hence, we shall now discuss the essential factors to remember

in employing a proficiency test of speaking skills.

One undeniable fact about communication is that any communication act takes place in a certain social context. Without understanding this social context fully the sender wishing to convey a particular message could fail in doing so, thus making his intended meaning totally unintelligible to the receiver. The point to be remembered here is that this could happen no matter how grammatically correct the sentence which the sender employed in sending the message are. Thus, in judging error gravity we may have to change our conventional notion of grammatical correctness into the appropriateness of the expression. For example, if a man from abroad visiting America uttered "Excuse me" unintentionally when he mistakenly went into the women's room, he may harshly be denounced by the women in the room, if he did not mean to be impolite at all. This suggests that the understanding of lexical meaning (to be found in a dictionary) of the two expressions ("Excuse me." and "I'm sorry.") is not enough, but that the man might not have been denounced if he had practiced the two expressions in several different contexts where one or the other expression is the only suitable one. Thus, in my view the appropriateness of an expression in a context should precede to the grammaticalness of an utterance in priority of evaluation. It should be only after this decision that judgement of grammaticalness comes in. Then, there arises the problem of judging how serious the error of any one utterance is in a scale of grammatical correctness. This should also be judged according to the degree of intelligibility of an intended utterance. However, very little research work has been done, so far, on this topic. An experiment of interest in this regard is reported by James (1977). In the experiment, a sample of fifty errors of written work was randomly selected from

24

learners of many origins (all of the errors were recognizable as error in no further context than the sentence containing it). The scoring was made on a basis of relative degree of gravity for each error by each team of twenty native and non-native speakers of English. There were several findings of significance in this study.　First, non-native judges tended to mark more severely than native speakers did. Relevant questions arise from this result; is there any positive correlation between their level of proficiency in English and the degree of severity in scoring? Do they become more tolerant of errors as they increase proficiency? The error types which native judges graded more gravely than non-natives and those which were scored more severely by non-natives than natives were also revealed. James puts it, "This confirms Richard's (1971) belief that native speakers are fussy about verb morphology, and suggests that native speakers tolerate lexical errors more readily than non-natives do. But perhaps native assistants should be asked to be a little less tolerant of these."[3] In order to justify the above statement, an adequate account of why native judges should be requested to be more critical of lexical errors must be made. Despite its difficulty inherent in the issue itself, a brief discussion will find its value.

　　If it is communicative competence at which the final goal of teaching expressive skills aim, we should seek to find a measure which roughly indicates the degree of seriousness of errors. The basic notion to be employed in setting up such a measure would be the one which considers to what extent an expressed sentence of oral or written performance causes the hearer or reader to misunderstand the intended meaning(s).　Hopefully such a measure, when construct-ed, will tell us the validity of a previously cited statement of James. But I suspect that the statement is a little hasty　one considering the

fact that it is the lexicon that is the most difficult area to overcome even for upper-intermediate and advanced learners, not to mention of the beginners. Consequently, many misunderstandings can take place owing to the wrong choice of the words in the context. If this assumption is correct, then it becomes imperative that we establish a set of standards to be able to score the learner's work of expressive skills. Of course, the prerequisite for the establishing of such standards is that we need to agree on the criteria on which the standards are based. This is probably a much more significant task to be completed ; hence, it is worth investigating more vigorously.

A pioneering work on this issue was done by Burt (1975). According to Burt, successful communication takes place when both the speaker and writer get across what he had in mind to the listener or reader. She conducted research from the listener or the reader's point of view ; if the receiver of the conveyed message does not correctly comprehend, successful communication will not result. One of the major purposes of Burt's work was to rank in order the kinds of linguistic categories which included most serious errors as felt by native and non-native judges. Such ranking of linguistic categories in order of seriousness was attained by multiplying the number of errors committed by the error gravity (represented by a 0-5 scale). She collected thousands of recorded tapes of spontaneous conversations and written works of students (learning English as a foreign language) such as compositions and letters. These materials containing two or more ungrammatical errors were then scored by native English speakers (the company janitor, the car mechanic, and the shopkeepers) on a basis of the relative comprehensibility. Each native speaker was presented four sentences, each of which included a few errors and asked which of these was the easiest to comprehend.

A few examples will illustrate the point.[4] The original sentence which the learner produced :

English language use much people.

For this original sentence each correction was made in the following manner.

1. The English language use much people. (the inserted)
2. English language use many people. (much corrected)
3. Much people use English language. (word order corrected)
4. The English language use many people. (the and much corrected)

While all these five sentences contain two or more errors, sentence three was unanimously considered as the most comprehensible, indicating that the word order was the major reason to make sentence three the most intelligible. Another example will follow :

The original sentence :

Not take this bus we late for school.

1. We not take this bus we late for school (we inserted)
2. Do not take this bus we late for school (do inserted)
3. not take this bus we will late for school (will inserted)
4. not take this bus we be late for school (be inserted)
5. If not take this bus we late for school (if inserted)
6. We do not take this bus we will be late for school (four omitted words in 1-4 inserted)

The five sentences (1, 2, 3, 4, 6), though partially or mostly corrected, were incomprehensible for all the judges. They even suggested the possibility, though unlikely, of the first four sentences to have the meaning : "We shouldn't take this bus. If we do, we'll be late for school." The crucial element to make the intended meaning clear, this time, was the insertion of "if." It is significant to note that sentence six, with the four corrections made, was still ambiguous. Burt foll-

27

owed this procedure with about 300 sentences containing more than one error and found in them that there exist certain types to hinder communication and others which do not impede communication. She labeled the former as "global" and the latter "local." Some of the categories included in the former type are :

    a. Wrong word order

       example : English language use many people.

    b. Missing, wrong, or misplaced sentence connectors

       example : (if) not take this bus, we late for school

The liguistic elements to be labeled as "local" were : errors in noun and verb inflections, articles, auxiliaries, and the formation of quantifiers. The term "local" was assigned because local errors were limited to a single part of the sentence, not affecting the comprehension of the whole sentence. Other grammatical structures which cause "global" errors are, according to her, psychological predicates complement system : verbs such as delight, thrill, charm, excite, impress, please, etc. and adjectives such as good, wonderful, important, necessary, easy, etc. When these groups of verbs or adjectives are used to express psychological states or reactions toward someone or something, the reverse order of experiencer and stimulus takes place (the experiencer is one who experiences and the stimulus is the thing or person that makes the feeling come out). A few examples will make the point clear :

    1. He doesn't bother the cat.

    2. I don't amuse that.

    3. He doesn't interest that group.

In the complement type of errors the learner often fails to insert the infinitival subject as in : (1) We want (him) to go to New York next week. (2) I couldn't walk yet after the baby was born so the doctor

didn't want (me) to go home. (3) Mother has a lot of work. Daddy expects (her) to stay at her office late. Often "global" errors result when the learner fails to know that certain verbs require the subject to be inserted in both the main clause and in the subordinate clause even if it is the same in both clauses. The following is a point in case :

    1. Anna told the priest to have six children.

       Intended : Anna told the priest that she had six children.

    2. He found out to be healthy.

       Intended : He found out that he was healthy.

The verbs which cause this type of errors are : think, know, find out, report, tell, notice, etc.

It is needless to say that there are numerous types of "global" errors which prevent communication. In this study "global" errors and "local" errors were assigned when syntactic errors were made. Depending on the context, a wrong choice of a single word whether it is a noun, adjective, adverb, conjunction or even article may well cause "global" errors. These errors which come from a wrong choice in the context may come into the category of lexicon, but there may be cases in which the errors are made in between the syntactic and lexical category. Further study in this area needs to be pursued. The corraboration of native speakers in judging the gravity of errors will be necessary, for they usually are the receivers to whom the message is transmitted. At any rate, there is no doubt that this type of study will reshape and improve the notion of proficiency in expressive skills, which is no doubt much more difficult to construct, score, and evaluate than are the receptive skills.

Implications

We have so far been discussing the key notions on which the test

of oral proficiency is based. In an attempt to elicit one's oral proficiency several other issues need to be dealt with.

One such issue would be that a different theoretical framework needs to be developed for different proficiency levels : the beginning, the intermediate and the advanced. Such a theoretical framework should be one which fairly clearly differentiates each proficiency level. It is natural that as one proceeds to acquire proficiency, he is asked inevitably to respond properly to the other party's response, thus requiring him to be equipped with competence in listening skills. The problem here is how we should evaluate this competence. The present writer feels that at least for intermediate and advanced level students this reciprocal ability of speaking properly in relation to the other's response should also be measured.

Another serious task imposed on us would be to search for an effective measure which will lead to uncovering the differences in oral proficiency, in other words how we know that one person knows more expressions than another. One solution may be to impose "time constraint" on test taking. Although we have no assurance that this factor alone will reveal the data implying the individual differences, there is evidence showing that when a testee is given ample time he could communicate somehow to the other party. Experimental data and our common sense tell us that when the time to express ideas is limited, we unconsciously try to select one of the expressions we know, and this is done in a flash of moment. When one is asked to respond orally, the more varieties of expressions one knows the more chances he has in responding quickly. This is only because if one knows only one expression and forgets that expression for some reason he has no way to resort to other expressions. It is thus hoped that with "time constraint" imposed on the testee and with a fairly

large number of test items, individual differences in oral proficiency will emerge. More researah work in other aspects may be needed in search for effective measures to pinpoint individual differences in oral proficiency.

NOTES

1. Jean-Yves Dommergues and Harlan Lane, "On ·Two Independent Sources of Error in Learning the Syntax of a Second Language." *Language Learning*, 26 (June, 1976), 115.
2. *Idid.*, p. 117.
3. Carl James, "Judgements of Error Gravities." *English Language Teaching Journal*, XXXI (January, 1977), 122.
4. Marina K. Birt, "Error Analysis in the Adult EFL Classroom." *TESOL Quarterly*, 9 (March, 1975), 55-61.

REFERENCES

Briere, Eugene J. "Are We Really Measuring Proficiency with our Foreign Language Tests." *Foreign Language Annals,* 4 (May, 1971) 385-391.

Briere, Eugene J. "Current Trends in Second Language Testing." *Papers on Language Testing.* Edited by Leslie Palmer and Bernard Spolsky. Washington D.C. : Teachers of English to Speakers of Other Languages, 1975.

Burt, Marina K. "Error Analysis in the Adult EFL Classroom." *TESOL Quarterly*, 9 (March, 1975) 53-63.

Cowan, J. Ronayne. "Reading, Perceptual Strategies and Contrastive Analysis." *Language Learning*, 26 (June, 1976) 95-109.

Davies, Alan. ed. *Language Testing Symposium.* London : Oxford University Press, 1968.

Dommergues, Jean-Yves and Lane, Harlan. "On Two Independent Sources of Error in Learning the Syntax of a Second Language." *Language Learning*, 26 (JUne, 1976) 111-123.

Etherton, A.R.B. "Error Anaysis : Problems and Procedures." *English Language Teaching Journal*, 32 (October, 1977) 67-78.

Folland, David and Robertson, David. "Towards Objectivity in Group Oral

Testing." *English Language Teaching Journal*, 30 (January, 1976) 156-167.

Gaies, Stephen J.; Gradman, Harry L.; and Spolsky, Bernard. Toward the Measurement of Functional Proficiency: Contextualization of the Noise Test." *TESOL Quarterly*, 11 (March, 1977) 51-57.

Hwang, Kwo-Yann and Dizney, Henry F. "Predictive Validity of the Test of English as a Foreign Language for Chinese Graduate Students at an American University." *Educational and Psychological Measurement*, 30 (Summer, 1970) 475-477.

Ingram, Elisabeth. "Testing in the Context of a Language Learning Experiment." *Language Learning*, Special Issue, 3 (August, 1968) 147-161.

James, Carl. "Judements of Error Gravities." *English Language Teaching Journal*, 31 (January, 1977) 116-124.

Johansson, Stig. "The Uses of Error Analysis and Contrastive Analysis (1)." *English Language Teaching Journal*, 29 (April, 1975) 246-253.
"The Uses of Error Analysis and Contrastive Analysis (2)." *English Language Teaching Journal*, 29 (July, 1975) 330-336.

Levenston, E.A. "Aspects of Testing the Oral Proficiency of Adult Immigrants to Canada." *Papers on Language Testing*. Edited by Leslie Palmer and Bernard Spolsky. Washington D.C.: Teachers of English to Speakers of Other Languages, 1975.

Levine, Josie. "An Outline Proposal for Testing Communicative Competence." *English Language Teaching Journal*, 30 (January, 1976) 128. 135.

Oller, John W., Jr. "Discrete-Point Tests Versus Tests of Integrative Skills." *Focus on the Learner*. Edited by John W. Oller, Jr. and Jack C. Richards. Rowley, Massachusetts: Newbery House Publishers, Inc., 1973.

Palmer, Adrian S. "Testing Commucation." *IRAL*, 10 (February, 1972) 35-45.

Ross, Janet. "The Habit of Perception in Foreign Language Learning: Insights into Error from Contrastive Analysis." *TESOL Quarterly*, 10 (June, 1976) 169-175.

Schachter, Jacquelyn and Celce-Mucia, Marianne. "Some Reservations Concerning Error Analysis." *TESOL Quarterly*, 11 (December, 1977) 441-451.

Spolsky, Bernard; Murphy, Penny; Holm, Wayne; and Ferrel, Allen. "Three Functional Tests of Oral Proficiency." *Papers on Language Testing*. Edited by Leslie Palmer and Bernard Spolsky. Washington D.C.: Teachers of English to Speakers of Other Languages, 1975.

Spolsky, Bernard. "Language Testing—The Problem of Validation." *TESOL Quarterly,* 2 (June, 1968) 88-94.

Spolsky, Bernard. "What Does It Mean to Know a Language, or How Do You Get Someone to Perform His Competence?" *Focus on the Learner.* Edited by John W. Oller Jr. and Jack C. Richards. Rowley, Massachusetts: Newbury House Publishers, Inc., 1973.

Tailor, Barry. "The Use of Overgeneralization and Transfer Learning Strategies by Elementary and Intermediate Students in ESL." *Language Learning,* 25 (June, 1975) 73-107.

Upshur, J.A. "Objective Evaluation of Oral Proficiency in The ESOL Classroom." *TESOL Quarterly,* 5 (March, 1971) 47-59.

Upshur, J.A. "Productive Communication Testing: Progress Report." *Focus on the Learner.* Edited by John W. Oller, Jr. and Jack C. Richards. Rowley, Massachusetts: Newbury House Publishers, Inc., 1973.

Yarmohammadi, Lotfollah. "Universal Versus Language Specific Tests." *English Language Teaching Journal,* 29 (October, 1974) 65-68.