

Developing Oral Placement Tests for Community-Based Language Programs

Sheila Dermer Applebaum and Elizabeth Taborek

This paper will describe the steps taken in developing oral placement tests for adults in two agency-based community ESL programs. A general overview of recent oral testing approaches will be

presented and the suitability of specific tests for use in meeting the expressed needs of community agencies will be discussed.

The Toronto Board of Education provides over 500 adult ESL classes, some of which are co-sponsored by community agencies. Working in cooperation with the community agency's ESL coordinator whose salary is funded by the Ontario Ministry of Citizenship and Culture, the Board sets up ESL programs, gives curriculum and administrative direction, hires and pays for instructors, and provides supplies.

Last year several agencies expressed a need for help with establishing testing procedures, initially for placement purposes and later for measuring student achievement and evaluating ESL programs. Up to that point, learners had been interviewed by the agency's ESL coordinator and placed at a particular level. The placement decision was based on the coordinator's familiarity with the program. Although this non-threatening placement procedure was sufficient to meet certain agency needs, in general it was not related to the curriculum of a program in any organized fashion. In addition, the interview could not provide data for achievement or evaluation purposes at a later date, nor could it clearly differentiate among the proficiency levels of the learners who tended to cluster at the lower end of the Educational Testing Service Scale (-0+ to 1) (Liskin-Gasparro 1982).

The process of developing placement procedures was begun with two different community agencies, Woodgreen Community Centre in the east end of Toronto and the Chinese Interpreter and Information Services (CIIS) in the west end. Both centres serve a large Chinese immigrant population and offer five levels of ESL classes, four of which are taught in a bilingual setting. A bilingual ESL program differs from regular ESL programs by providing an instructor who is fluent in the native language of a particular student group. The first language is used to

facilitate this understanding of difficult concepts in English at the beginner level.

After considerable discussion, it was decided to continue using an oral interview test. The following summarizes a number of different oral testing models examined in the process of developing a satisfactory instrument for the community agencies.

REVIEW OF ORAL TESTING

The Oral Interview

The oral interview is a performance test which is communicative because it approximates real-life tasks that individuals are frequently called on to perform in society such as answering biographical questions. The oral interview has the advantage of high face validity, as it sets up a natural interactive speaking situation between interviewer and interviewee. In addition, it has high content validity, provided the testing tasks are closely related to tasks taught in the curriculum at different levels. Reliability or stability of test scores is possible if examiners are trained in giving the oral interview test and in overcoming subjectivity in scoring. There is one possible drawback, however. The practicality of this kind of test can be a problem because of its costs in time and in staffing.

Types of Oral Interviews

Foreign Service Oral Interview (FSI)

The best known measure of oral proficiency is the Foreign Service Oral Interview (FSI). The Educational Testing Service Manual (Liskin-Gasparro, 1982) is an excellent reference for the procedure for this test which has been refined and researched over the years. The purposes of this interview are different from what would be expected in community-based programs and the test itself may be too academic in nature. However, there are some areas that are adaptable. Particularly appropriate are the phases of warm-up, level-check, probe, and wind-down (Liskin-Gasparro, 1982, p. 62).

The Warm-Up

Greetings and small talk help to put candidates at ease. Remarks that are easy for the candidate to respond to will do much to set the tone for the interview.

The Level-Check

The objective of this phase is to determine the highest level of oral

proficiency the candidate can reach in continuous speech. It is necessary to check a number of topic areas and tasks to be sure. When a candidate appears comfortable speaking at a certain proficiency level and uses language which has appropriate content and accuracy, then the tester can probe to a higher level.

Probe

To probe, the tester asks questions on topics which require a higher level than has been decided upon at the level check. Several probes should be tried. If there is a sharp drop in grammatical accuracy, correct vocabulary or fluency, the tester should return to an easier level.

Wind-Down

The objective here is to leave candidates with a feeling of satisfaction after having been pushed to the limits of their spoken ability. The tester should then return to the level where candidates were most at ease.

This test provides useful descriptions of the various levels and of scoring criteria. Two raters are recommended rather than only one in order to arrive at a more objective scoring of the interview.

The Ilyin Oral Interview

Donna Ilyin (1976) designed an oral interview kit for immigrants who may not be able to read or write. It has several good points. Only one examiner is needed and the testing time can be relatively short depending on a student's proficiency level. It has a high inter/intra reliability factor (i.e. reliability between different raters and different items). It uses pictures about the daily life of a student as cues. The examiner asks questions about the pictures and scores the candidate on the answers. One disadvantage is that the scoring is based on structural accuracy rather than communicative success. Moreover, questions do not approximate natural speech. In one example, a candidate is required to ask a question about a picture using "if". It was decided that this test was not suitable for a bilingual program because the lower proficiency level learners would not even be able to get through the pre-test. As a result, the Ilyin test could not be used to distinguish among students within the lower levels of the programs.

Object-Oriented Testing

In his article on object-oriented language testing, Gordon (1974) noted that it was difficult to get candidates to talk naturally. He suggested the use of objects as a means of getting the candidates to describe or compare. He used a number of related books and asked

questions about them. Critics of the test have warned, however, that while props may be appropriate for natural discussion, objects must be relevant to the candidate's own experience. The test must also avoid making adults feel that they are being treated like children in elementary school.

Group Testing

In their work on group testing, Folland and Robertson (1976) had up to seven students listen to a tape and then discuss the topic it referred to. The tape was based on classroom sessions and was similar to the type of lectures they heard and discussed every week. The authors felt that the language in the test was realistic, that the students were more at ease in a group than they were by themselves and that the cost of such a test was minimal. It should be pointed out, however, that those administering the test must be clear about whether they are testing speaking, or listening and note-taking skills.

Performance in Authentic Situations

Hinofotis (1981) described an appropriate real-life task used as an oral test to decide on the hiring of teaching assistants whose mother tongue was not English. The candidates were asked to explain a difficult term from their field to a student who did not understand it. This interaction was videotaped and several examiners scored the performance, not only on language proficiency, but on the success of the communication, eye contact, delivery, confidence, and other factors. This test assessed performance in a situation similar to those occurring in the candidate's work.

OTESL

In the work being done on the OTESL (Ontario Test of English as a Second Language in Colleges and Universities) (Mendelsohn & Tumpene, 1985), the writers first examined what students entering university or college could be expected to do in English. Their final list included: ask and answer questions; express an opinion; support an opinion; rephrase; express a counter-argument. It was only after these needs had been identified that the researchers set out oral tasks to match. Thus, the kinds of questions in the interview were related to the type of speaking tasks that a university student would be expected to perform. In one section, for example, the candidate is asked to rephrase material he/she has just read—a task that is necessary in university/college courses.

George Brown College Oral ESL Assessment

The ESL staff at George Brown College in Toronto has also created an oral test which reflects communicative teaching principles. (See Appendix 2.) The test gives information to the student in advance and lets him/her choose topics to talk about. At the intermediate level, for example, students are asked to choose two topics to compare from a list of suggestions—a task they have already done in class. Scoring consists of an overall rating based on a series of bands which describe proficiency levels such as moderate speaker, limited speaker, etc. Candidates are given copies of the completed assessment with comments, thus providing them with immediate feedback.

CRITERIA FOR TESTING WITHIN COMMUNICATIVE LANGUAGE PROGRAMS

Oral testing in the 1980s has become communicative, reflecting the communicative approach to language teaching. From the preceding review of recent oral testing approaches it can be seen that there are specific criteria which, if included, should lead to successful oral testing in a program which is taught communicatively. First of all, an oral placement test should test tasks and topics covered in the program's existing curriculum. It should use language that is as natural as possible. The candidate should be put at ease and involved as much as possible in the test itself. For example, candidates can be told ahead of time what the questions and topics will be. They can even be asked what proficiency level they think they fit into. Group oral testing is recommended in order to save time and create a more congenial and supportive testing climate. Oral tests can also be used for achievement rather than for placement purposes, but these should take place frequently and data from them should be collected over time. The instructor's evaluation of the student's level should also be considered as another aspect of evaluation. The most important factor to note is that communicative teaching should be evaluated through communicative testing.

The general principles of communicative language testing which guided Swain and her colleagues (1982) in their large-scale achievement testing study provided a useful framework for the development of the pilot community-based testing programs.

Start From Somewhere

It was first necessary to look at what was already in place within the two community-based ESL programs: current registration forms; proce-

dures for placement; typical tasks being taught within each level of the actual curriculum; curriculum booklets for each level. In addition, a survey was designed and administered in order to focus on the specific testing needs of instructors and to involve them more directly in the test development process. For example, at the Chinese Interpreter and Information Services, the current registration form was modified and incorporated as part of the placement test. (See Appendix 1.)

Concentrate on Content

Having instructors focus on the *content* of typical tasks and topics covered in each level was particularly difficult, although rewarding. (See Appendix 2.) There was a lack of agreement as to what constituted a communicative speaking task. There were difficulties in describing tasks which incorporated the four components of communicative competence (Canale & Swain, 1980) including grammatical, discourse, and sociolinguistic competence along with a knowledge of their own learning strategies. In addition, an analysis of tasks revealed discrepancies (gaps and overlapping) across classes taught at the same level.

Concentrating on content *did* produce positive results: an increase in instructor awareness of the speaking and writing tasks actually used in class; an accurate description of the range of skills expected at various proficiency levels; an agreement on typical tasks for each level; an agreement on the validity of current levels and on the possibility of adaptation. It became evident that by examining a program in this manner, an agency could provide a clearer description of its objectives and could later evaluate whether these were being met.

Bias for Best

The principle of "doing everything possible to elicit the learner's best performance" (Swain, 1982) reinforces the agency's concern for its clients and the tester's desire to assess with confidence what a learner can do in English if given the opportunity. For example, the learner's performance can be facilitated by allowing ample time to fill in the registration form, by providing clear, non-threatening instructions about the purpose and format of the test (in their native language if requested), and by giving permission to ask questions and to correct errors.

Work for Washback

Because "washback" refers to the effect that a test has on the choice of curriculum items a teacher uses in the classroom, this fourth principle could only apply to proficiency tests given during the ESL course or to placement tests given to students continuing on at the end of the year.

Since the data collected for the initial placement test were based on a description of the curriculum of the current ESL program, the same information could be used again. (See Appendix 2.) This process would allow for a maximum "washback" effect. In addition, involving the instructors and other ESL personnel in the development, administration and scoring of the tests helped everyone to agree on various aspects of the test. It should be noted that the questionnaire that instructors completed underlined their ambivalence about telling learners what to expect during the testing procedures. However, it would be beneficial to show the instructors how their own comfort with test development and change in the program was heightened by being treated as an adult in the consultation process and by having useful information about what to expect from the tests. These principles were referred to throughout the project and influenced the type of scoring and criteria used to evaluate learners.

Criteria for Evaluation

The oral assessment form and scoring criteria (see Appendix 3) were written with the lower proficiency levels in mind. Levels of proficiency were described so that there would be more differentiation at the lower end of the scale where many community ESL programs have the largest number of clients. The descriptions have been adapted from Carroll's bands (1980), George Brown College ESL Oral Assessment Form (1985) and Liskin-Gasparro (1982). The performance factors were chosen after several trials with various criteria from other tests. A brief explanation of the performance factors follows:

Fluency

The tester should observe how smoothly the candidate's speech flows and should note the number of awkward pauses that interfere with comprehensibility. Fillers should be natural and not an interference to communication.

Pronunciation

The rater should observe errors in the production of English speech sounds. The rater should also consider the degree to which stress and intonation patterns imitate the target language and do not interfere with comprehensibility.

Linguistic Skills

The rater should observe the candidate's control of the general vocabulary which allows for successful communication on a certain topic. The rater should also observe structural errors and note how much they either interfere with communication or distract the listener.

Communicative Success

The rater should observe the candidate's use of language strategies to prevent communication breakdown when other skills are lacking. Communicative success should also include the ability to initiate and maintain conversation, take turns and use conversational gambits.

Listening Comprehension

The rater should observe the ability of the candidate to comprehend the questions and conversation of the examiner.

Current Difficulties in Oral Testing Implementation

Lack of funding for the administration of placement tests is considered the most significant obstacle to implementation. If registration for classes is high, more time and testers are required for quality testing. Consequently, more salary dollars are needed. To achieve stability in test scores, it is necessary to provide funding for the training of instructors in the area of testing theory and procedures.

Another concern is the anxiety created by the negative image of testing and the normal reaction to any change in procedures that both community agency personnel and ESL students exhibited. Teachers and ESL coordinators could help solve the problem of the threatening image of tests by giving test information to their clients before the placement assessment. This could be done in the first language. Teacher education courses could offer a session on communicative testing and workshops could supplement initial training. As an expansion of this pilot placement test, instructors might be encouraged to examine its applicability to other areas within bilingual ESL programs. For example, it could be adapted for achievement evaluation and English in the Workplace (EWP) placement of students.

Future Directions

In the ongoing work on placement procedures, plans are underway to produce a videotape reflecting the principles and procedures discussed in this paper. Students will be interviewed and then instructors will be asked to view the videotapes and make judgments on the proficiency levels of the interviewees according to scales which have been discussed in in-service sessions. These suggestions form a model upon which to build and adapt testing to suit particular administrative and program needs. Further work must be done to make this test a practical and useful addition to community programs.

REFERENCES

- Canale, M. (1981). Communication: How to evaluate it? *Bulletin of the Canadian Association of Applied Linguistics*, 3 (2), 77-94.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1 (1), 1-4-7.
- Carroll, B. (1980). *Testing communicative performance*. Oxford: Pergamon.
- Clark, J.L.D. (1979a). Psychometric considerations in language testing. In B. Spolsky, (Ed.), *Advances in language testing series: 2*. Arlington, Va.: Centre for Applied Linguistics.
- Clark, J.L.D. (1979b). Direct versus semi-direct test of speaking ability. In E.J. Briere & F.B. Hinofotis (Eds.), *Concepts in language testing: some recent studies*. Washington, D.C.: TESOL.
- Clifford, R. (1981). *Oral proficiency performance profile and global rating*. Paper presented at the Pre-conference on Oral Proficiency Assessment, Georgetown University Round Table, Washington, D.C. mimeo.
- Edmonton Board of Education, Continuing Education Department, (1982). *Adult ESL Curriculum Guide*.
- Engelskirchen, A., Cottrell, E., & Oller Jr., J. (1981). A study of the reliability and validity of the Ilyin Oral Interview. In A. Palmer, P. Groot and G. Trosper (Eds.), *The construct validation of tests of communicative competence*. Washington, D.C.: TESOL.
- ETS Oral Proficiency Manual*, (1982). Princeton, N.J.
- Folland, D., & Robertson, D. (1976). Towards objectivity in group testing. *English Language Teaching*, 30, 56-67.
- Galway, J., & Whittington, A. (1984). Performance tasks: An assessment technique used at TOSTP. *TESL Talk*, 15, Winter/Spring, 36-40.
- Gordon, W.T. (1974). Object-oriented oral testing. *The Canadian Modern Language Review*, 31, 72-74.
- Hinofotis, F., Bailey, K., & Stern, S. (1981). Assessing the oral proficiency of prospective foreign teaching assistants: instrument development. In A.S. Palmer, P.J.M. Groot, & G.A. Trosper (Eds.), *The construct validation of tests of communicative competence*. Washington, D.C.: TESOL.
- Ilyin, D. (1976). *Oral Interview Test*. Rowley, Mass.: Newbury House.
- Jones, R.L. (1977). Testing: a vital connection. In J.K. Phillips (Ed.), *The language connection: From the classroom to the world*. Skokie, Ill.: National Textbook Co.
- Jones, R.L., & Spolsky, B. (Eds.). (1975). *Testing language proficiency*. Arlington, Virginia: Centre for Applied Linguistics.
- Levenston, E.A. (1975). Aspects of testing the oral proficiency of adult immigrants to Canada. In L. Palmer & B. Spolsky, (Eds.), *Papers on language testing. 1967-74*. Washington, D.C.: TESOL.
- Liskin-Gasparro, J. (1982). *ETS Oral Proficiency Testing Manual*. Princeton, N.J.: ETS.
- Masden, H., & Jones, R. (1981). Classification of oral proficiency tests. In A. Palmer, P. Groot, & G. Trosper (Eds.), *The construct validation of tests of communicative competence*. Washington, D.C.: TESOL.

- Mendelsohn, D., & Tumpane, M. (1985). *OTESL oral test*. Personal Interview, Toronto, Ontario.
- Spolsky, B., Murphy, P., Holm, W., & Ferrel, A. (1975). Three functional tests of oral proficiency. In L. Palmer and B. Spolsky (Eds.), *Papers on language testing*. 1967-74. Washington, D.C.: TESOL.
- Swain, M. (1982). *Large scale communicative language testing: a case study*. Plenary paper presented at the International Symposium on Language Testing, University of Hong Kong, December. mimeo.
- Swain, M., Dumas, G., & Naiman, N., *Alternatives to spontaneous speech*. (EDRS: ED123872).
- Tyacke, M., & Mendelsohn, D. (1982). *Oral interview and oral interaction*. School of Continuing Studies, University of Toronto. mimeo.
- Upshur, J.A. (1975). Objective evaluation of oral proficiency in the ESOL classroom. In L. Palmer and B. Spolsky (Eds.), *Papers on language testing*. 1967-74. Washington, D.C.: TESOL.

THE AUTHORS

Elizabeth Taborek and Sheila Dermer Applebaum are Administrative Lead Instructors for Adult ESL Programs in the Continuing Education Department at the Toronto Board of Education. Both have been responsible for teacher-training, materials development and the supervision of programs.

- LEVEL 2 (Basic/Low; bilingual)
- to follow directions—community map
 - to make an appointment
- LEVEL 3 (Basic/High; bilingual)
- to make requests
 - to leave telephone messages
- LEVEL 4 (Intermediate/Low; bilingual)
- to express likes and dislikes
 - to relate a story
- LEVEL 5 (Intermediate/High; unilingual)
- to make a statement and provide support
 - to compare and discuss present and past experiences

APPENDIX 2
ESL CURRICULUM GUIDE
WOODGREEN COMMUNITY CENTRE
LEVEL 1

OVERALL OBJECTIVE

Learners should have minimal competence using the English writing system. Learners can express simple elementary needs and can operate in a very limited capacity in the community. (Adapted from ETS/82.)

PERFORMANCE OBJECTIVES

Upon completion of this program, the learner should be able to:

A. Literacy

- read and write numbers, identify and write letters, spell letter names, read basic sight words related to his/her survival needs, write basic numbers and words used in filling out forms, associate spoken word/sentence with written word/sentence, tell time and use the calendar, use Canadian money, use a simple street map.

B. Oral

SOCIALIZING

- introduce self and greet new people (simple form), give personal information, use common polite questions.

BASIC COPING SKILLS

- give and follow simple instructions, identify and describe common objects, identify and describe people, make simple requests, describe daily routine.

SHOPPING

- use money when shopping for food or clothing, identify common foods.

TRANSPORTATION

- use a simple street map to find a specific location, use the transit system, request and give simple directions for getting somewhere.

TELEPHONE

- request emergency assistance.

STRUCTURAL COMPONENT

Students should learn the grammatical structures they need to express the functions. Achieving fluency and confidence with these functions is the priority. Refining their language for grammatical accuracy follows. (A New Start - Canada—Selman)

Structures

- simple present tense—affirmative, negative
- to be, present affirmative, negative
- to have, present, affirmative, negative

- imperative (pick up, open, turn left, etc.)
- prepositions of location
- plural forms of nouns
- subject/verb agreement

RECOMMENDED TEXTS

A New Start - Canada, M. Selman
 Lifelines I, B. Foley
 Before Book One, Boyd & Boyd

**ESL CURRICULUM GUIDE
 CHINESE INTERPRETER & INFORMATION SERVICES**

LEVEL 2

- DESCRIPTION:** Basic/Low; bilingual
- FUNCTION:** Communicates with memorized material.
 Answers simple questions.
 Writes simple information.
- CONTENT:** socializing (introductions, minimum courtesy requirements), time, everyday survival topics (shopping, banking, transportation, health, post office, community services, telephone, eating out, holidays, housing, work)
- ACCURACY:** intelligible to a native speaker used to dealing with immigrants; makes frequent errors that interfere with communication

Upon completion of this program, the learner should be able to meet the following performance objectives:

TASKS	TOPICS
to follow directions (community map, T.T.C.)	transportation
to understand instructions (e.g. class)	survival (coping) skills
to give simple directions/location of familiar places/streets	T.T.C./travelling
to identify and express weekdays, months, year, today's date, time	dates/time
to introduce oneself, family, friends, colleagues	socializing
to greet and leave people properly	socializing
to ask simple questions	shopping, community services
to use the telephone to ask for a person	survival (coping) skills
to respond to a "wrong number"	survival (coping) skills
to make a telephone call for emergency	survival (coping) skills

to say your phone number/social insurance number/to count and use money
to be aware of simple banking procedures (e.g. deposit, withdrawal, writing a cheque)
to identify and read bills
to use minimum courtesy requirements
to order food in a fast-food restaurant
to read a simple menu from a fast-food restaurant
to carry out simple postal procedures
to make an appointment
to identify parts of the body
to identify common food, household, clothing items
to read labels
to read signs
to describe common illnesses
to describe common occupations

survival (coping) skills
survival (coping) skills
survival (coping) skills
socializing
survival (coping) skills
survival (coping) skills
survival (coping) skills
survival (coping) skills
survival (coping) skills
survival (coping) skills
survival (coping) skills
survival (coping) skills
survival (coping) skills

APPENDIX 3

DRAFT COMMUNITY BILINGUAL ADULT ESL PROGRAM ORAL ASSESSMENT SCALE*

<u>ETS</u>	<u>BAND</u>	<u>EQUIVALENT TO NATIVE SPEAKER</u>
5	9	
		<u>ADVANCED SPEAKER</u>
4	8	Fluent. Seldom gropes for appropriate vocabulary. Can use some idiomatic expressions. Makes few errors in grammar. Can discuss abstract topics. Can hypothesize, persuade, argue. Can understand most formal discourse within the range of his/her experience.
		<u>PROFESSIONAL COMPETENCE</u>
3	7	Can discuss opinions in informal situations on political, social and professional problems. Good control of grammar though occasional errors in complex constructions. Broad vocabulary. Pronunciation does not interfere with comprehension by native speaker. Can understand almost all native speaker conversation at normal rate of speech and can understand some formal discourse (e.g. formal speech, news) if not too long or dense.
		<u>COMPETENT SPEAKER</u>
2	6	Has more flexibility and range. Can initiate change in conversation. Shows more independence in discussion. Makes some linguistic errors in low frequency structures. Uses past and future tenses easily. Can participate in conversation about current events with native speaker.
		<u>GOOD SPEAKER</u>
1+	5	Foreign intonation evident. Hesitant in speech. Maintains dialogue and follows topic switches. Can join sentences in limited discourse. Can understand attitudes and reactions. Control of frequently used patterns. Can talk on social topics and concrete areas such as educational background and work. Can describe, give instructions, report and discuss past events.
		<u>ADEQUATE SPEAKER</u>
1	4	Shows some initiative in keeping conversation going. Linguistic errors still frequent but shows some control of simple patterns. Can follow conversation but still needs some repetition. Vocabulary permits discussion on a few personal topics other than basic needs. Can create with the language.

* Adapted from ETS (1982) and George Brown College (1985).

		<u>MODEST SPEAKER</u>
Novice	3	Emerging signs of spontaneity and flexibility. Slight increase in length of conversation. Can ask and answer simple questions although inaccurate. Still makes many linguistic errors. Gets the gist of the conversation but still not master of it. Can expand on more topics related to basic survival needs. Intelligible to native speaker who is used to dealing with foreign speakers.
High		
<u>ETS</u>	<u>BAND</u>	<u>MARGINAL SPEAKER</u>
Novice	2	Can talk on basic survival topics such as weather and time. Can just hang on to the gist of conversation. Requires repetition or explanation. Makes errors that cause misunderstanding. Conversation hard to maintain. Answers to questions usually rely on memorized material. Difficult to understand.
Mid		
		<u>EXTREMELY LIMITED</u>
Novice	1	Operates in a very limited capacity on a few basic topics of personal need. Does not understand the gist of the conversation. Uses minimal response most of the time. Many hesitations and misunderstandings. Many linguistic errors. Extremely limited verb forms, grammar and vocabulary.
Low		
		<u>NO FACILITY</u>
	0	Learner cannot function in the spoken language. Knows and recognizes a few words or memorized phrases. No real conversation.