

Testing Should Reflect Teaching¹

David J. Mendelsohn

How many of you are practising classroom teachers? Therefore I would assume that the same number are involved in classroom language testing of one sort or another.

But how many of you switch off when you read or hear terms like *Regression*, *Analysis of Variance*, and *Biserial Correlation*? To say nothing of my favourite *Homoscedasticity*—defined as follows:

This is a property of bivariate distribution when the variance is the same (i.e. homogeneous) at all points along the regression line. It is also an assumption of analysis of covariance that all within-cell slopes are equal (i.e. homogeneous). (Henning 1987: 192).

This is *not* what this paper is about, and what is more, it is not being delivered by an expert in language testing who is going to blow you away with jargon, statistical terms and illegible overheads of tables of figures.

This paper is being delivered by a practising classroom teacher committed to teaching language communicatively, and committed to testing what he teaches—hence the title of the paper: *Testing Should Reflect Teaching*. I shall attempt in this paper to do a number of things:

- to spell out the essential philosophy of testing within a communicative framework
- to examine some common learner-attitudes to tests
- to discuss five principles I believe we should follow in designing second language tests
- to consider the question of how to test communicatively.

In short, my goal today is to remove some of the mystical aura that surrounds language testing and to suggest that classroom teachers are perfectly capable of constructing good reliable and valid tests, and what is more, that their input is imperative in all test construction undertaken by “experts”.

Before going into any further detail about testing of different kinds, I would like to emphasize the importance of the participation of the classroom practitioner along with the testing experts in the designing of large-scale tests. Sometimes test-development teams develop a test without this participation, and what often results is a very well-designed test that is in accordance with all the necessary testing principles, but one that the class-

room teachers do not like for one reason or another. This results in resistance to the test, a failure to exploit the pedagogical value of the test, and a handcuffing of the teachers—if their students are going to be taking this test that they don't like, they are still morally obliged to teach to it!

Definition of a Test, and the Essential Philosophy of Testing within a Communicative Framework

Shohamy (1985:3) defines a test² as:

A procedure or device for measuring and evaluating a person's language knowledge, based on current definitions of language. A test is a *sample* of that knowledge and needs to be a good representation of it.

I do not think it would be unfair to say that the goal of testing in the past was to find out what the learner knew *about* language. And we all know many examples of locally produced and commercially marketed tests that fall into that category. I believe that the goal of testing today, in keeping with my commitment to communicative language teaching, is to see what someone *can do with* the language.

So, if a test is attempting to assess a learner's communicative ability, it should be concerned with the learner's ability:

- (i) to produce and comprehend authentic language in context
- (ii) to participate in simulations of real life situations
- (iii) to handle, receptively and productively, larger pieces of connected language, not just single sentences
- (iv) to recognize and use the language as it varies over different topics, settings, speakers, etc.

Only little attention should be given to testing isolated and discrete elements of grammar and vocabulary.

Common Learner Attitudes to Tests

Classroom teachers are acutely aware of the fact that their students, regardless of age, have very clearly defined opinions about the language teacher, the lessons, the materials used, and, most certainly, about the tests they are given. We often disregard these opinions for one of two main reasons: we often feel that the learner's opinion is "wrong" or misguided, and that "we know better"; and, on the other hand, we often know that the learner's opinion is in fact correct, but do not really know how to change and improve the things that they are critical of.

Shohamy (1985) carried out a study in Israel on what students think of tests, and I will quote some of her findings that have the greatest bearing on our topic.

She begins by reporting that 90% of the students answered that they did *not* feel that their tests reflected their language. The following are comments by some of the students questioned:

- I never learn anything from tests because the teacher never corrects the mistakes I make, so I end up at the same place where I was before I took the test, except now I also have a bad grade.
- I don't see the connection between the test and my knowledge, otherwise how can I explain the fact that I get good grades on English tests, but last week when I met an American, I couldn't say anything in English? How come we never speak on tests?
- I think that it is really strange that whenever I study hard, I don't get a good grade, but when I don't study at all, I happen to succeed. Does it say something about me or about the test?
- What I hate most is when the teacher does not tell us in advance what the test will cover. It seems that I am always studying the wrong things.
- The test becomes a punishment—"Since you don't know the material, you will have a test on Monday."
- In class we learn how to speak but in tests we are always required to write.

A close examination of these learners' comments shows that they are in fact calling for communicative language tests that are fair, and reflective of what they've been taught.

For the purposes of this brief paper, I must assume that the teaching is communicative, and will devote my attention to the testing.

Five Basic Principles We Should Follow in Designing Second Language Tests

1. We should update outmoded tests and outmoded thinking about tests.

Second language methodology is not static—it is continually being revised and modified in keeping with the most up-to-date ideas on second language acquisition. And in most second language programmes, the methodology has been updated. Communicative language teaching has supplanted audio-lingual/structural methodology in most programmes, but, unfortunately, the language testing has not kept up. In many places the teaching is now communicative but the testing is not. Much of the testing is still of lists of decontextualized linguistic items and the learner is scored for the number of correct responses given, usually in a multiple-choice format—this format having been chosen not for any other reason than that it is easily and objectively scored. The items tested are usually those that *are likely to cause difficulties*—this, after all was the underlying principle of Lado's in his 1961 book, *Language Testing*, written as an outgrowth of his work on Contrastive Analysis.

What is more, these decontextualized tests of small features of linguistic

proficiency often *pose* as something that they are not. For example: The learner hears on tape: *When are you going?* and has to choose in the answer book between:

- (1) To the library
- (2) Tomorrow
- (3) By subway
- (4) Because I'm tired

This is not, in my opinion, a test of listening comprehension, although that is what it has traditionally been called because the stimulus is *listened to*. Just because the learner *hears* a sentence hardly makes it a valid test of their ability to comprehend spoken English. Nor is the following an example of an oral test, although that is what it is called:

The interviewer says: *I am going to Montreal.*

The test-taker is required to formulate the question: *Where are you going?*

This, in my opinion, is not a test of oral proficiency, but a grammar test carried out in an oral mode.

Carroll (1980:9) sums up the problem with these traditional tests as follows:

Detaching test items from their communicative context is to risk finding little about the learner's behaviour which is not trivial; and merely multiplying the number of trivia is not going to solve the measurement problem.

Below, I will discuss communicative language tests, but a word is in place on modifications that can be made to existing tests that will make them more communicative.

First, in all skill areas, and particularly in oral testing, when scoring, note should be taken of the aspects of communicative competence beyond the linguistic level i.e. of the sociolinguistic and discourse features of the language and the learner's comprehension or mastery of these.

In cases where teachers are locked into traditional test formats, Wesche (1981: 568-9) proposes adding some sub-tests or items clearly based on a view of communicative competence:

One can ask listening and reading comprehension questions based on a global understanding of the meaning conveyed through language use in context, rather than pinpointing discrete grammatical points.

2. *We should beware of the "Genesis Syndrome"*

In the book of Genesis it says: "And God looked on his work and saw that it was good." Despite the fact that we do not have God's infallibility,

we tend to look on work that we have put a lot of time and effort into in the same way. The problem is not that we should look on our work and see that it is good, but rather a reluctance to look on our work and admit when it is *not* good. This is what I call the *Genesis Syndrome*. In language teaching in general, and test design in particular, it is imperative that we look very critically and objectively at tests we construct, then try them out, and discard those parts that do not work or are not feasible.

3. We should ensure that our tests are as valid as possible

The terms *reliability* and *validity* used in presentations on testing usually trigger that anti-jargon defence-mechanism of switching-off.

However, the following five criteria for a good test would be readily agreed on by most classroom teachers:

- (i) The test must measure what we intended it to measure. We should begin designing every test by spelling out the specifications of what we want to test. Having done this, we are then able to ask whether our test is sufficiently representative and comprehensive. For the classroom teacher this boils down to the central question: is this test a fair reflection of what has been taught? The test I referred to earlier in which test-takers had to make questions from statements is not, in my opinion, a test of the oral proficiency of the test-taker—it does not tell us about the learner's ability to communicate orally. Neither would a test that asks two small questions be an acceptable test, because it would not be comprehensible enough.
- (ii) The test must be felt by the test-takers to be fair and to be testing what they need to know. This criterion is one based on the subjective opinion of the test-takers, but is nevertheless very important. If test-takers do not find that a test meets their expectations, or they find that this kind of test is very strange to them, it can cause them to perform very badly.
- (iii) If the test is being given to predict how successful learners will be at using the language in the future, for example, in their studies, then it must be shown to do this. For example, if you are giving a test after training tour-guides, you should compare results on the test with how well these learners can guide tours in the second language subsequently.
- (iv) The test should yield similar results to other tried-and-true tests of the same aspects of language. If you have designed a new test, for example, of reading comprehension, then the results on your test should not be wildly different from those on a good and trusted reading test.
- (v) The tests should yield similar results when scored by different markers, and when re-scored by the same marker on a different day. If the

results do not prove to be similar, then you would have to question the confidence you can place in your test results.

When these five criteria are met, we have what is called a valid and reliable test. When they are not met, we have the Genesis Syndrome, and must be ready to modify our test accordingly, regardless of the time and effort already invested, and the time and effort required.

4. We should ensure that our tests have a positive washback effect on our teaching

There is no question that tests have a very important influence on teaching. This is most noticeable in situations in which the test that the learners will have to take is not constructed by the teacher. The most obvious example would be the T.O.E.F.L.—if the teacher is teaching a T.O.E.F.L. Preparation Course, then the shape and content of the T.O.E.F.L. test will determine the shape and content of teaching. After all, if a teacher takes on the task of preparing students for the T.O.E.F.L., then that is what they should do—it would be professionally irresponsible to take on the task, and then do solely communicative oral work in class. Similarly, if students are going to be given an integrative communicative test of the type that Swain (1984) describes, and that I will be referring to later, it would be equally irresponsible to do mainly discrete-feature grammar work in preparation.

In both cases, the test will influence the teaching. Therefore the nature of the test is very important—it *must* be in keeping with the teaching you believe you should be doing. When, for example, the communicative test described by Swain is going to be used, it has the effect of the teachers doing classroom activities of a similar nature in preparation. This positive effect of the test on the teaching is known as a *washback* effect. When a test causes teachers to spend their time on activities that the teacher would not have chosen to do, this negative effect is known as a *backwash* effect.

This influence of the test back to the teaching is extremely important, and has to be borne in mind by the classroom teacher all the time. As Carroll (1985: 78) puts it, you should “look for tests which vitalize teaching, not lay a dead hand on it.”

It might appear, at first glance, that this is not an issue when the classroom teacher is designing the test. But ironically, it is. The unfortunate chain of events leading to negative backwash can go as follows: the classroom teacher does not really know how to design a communicative test of what she/he is teaching. This results in a test that does *not* reflect the teaching because the teacher has produced a traditional grammar-based discrete-item test. This will, in time, have a negative backwash effect on the teaching.

5. *We should design our tests carefully, based on precise specifications*

All too often we put together classroom tests without sufficient care, and without really knowing what we should be doing. My first recommendation in this regard would be for teachers to read some books on the subject. I would *highly* recommend Shohamy's book which exists in draft form and should be published commercially very shortly. Cohen's book is also very useful.

The starting point should be defining the objectives. In classroom testing, the testing objectives are determined by the teaching objectives—by what has been taught. Or, to be more accurate, the testing objectives *should be* the teaching objectives. This in turn calls for the teacher to specify what she/he is teaching and why. But, as I stated earlier, for the purposes of this paper we are assuming that this was done.

It is imperative that we spell out the *specifications* for our test precisely. Only by doing that, will we test what we teach. If we do not do this, we will begin the process of test design at the end—by thinking of tasks and test items. And when we do that, the items rather than what we have taught will shape the test.

The objectives are usually defined in today's testing in behavioural terms—we specify what the learner is expected to be able to *do* with the language. For example, in an oral test, this might include such things as: to give directions; to explain the stages in a process; to describe how something works, etc. Spelling out the objectives in the first stage in the process.³

Having listed the objectives, a decision has to be made as to what material will be tested i.e. the content. This requires that the teacher identify the subject matter and topic to which the student is expected to respond for each objective. When selecting the content, every effort should be made to make the test interesting. This is not an objective in test design that is taken very seriously. But if we want the test to elicit our students' best efforts, then we should strive to make it interesting. In addition, the selection of the material should be made, bearing in mind that it should not make the students unduly anxious. A little anxiety is acceptable, but studies have shown that when students are overly anxious, this has a debilitating effect.

The third stage is to decide on the appropriate *weight* to be given each objective and content area. This should be determined by the time spent and the emphasis placed on this feature in teaching. This is another very important way in which testing should reflect teaching.

Working out these specifications provides a blue-print or plan that will result in sensible and fair selection of test items. I would not hire an architect or engineer who begins building before working out their plans.

I would not hire a surgeon who begins cutting before working out the plans. This would be too ludicrous and arbitrary to contemplate. It is my contention that to sit down to design a test, and to begin by deciding on test items is equally irresponsible. Only when the planning is complete, should the test items be chosen.

The selection of the items itself is a difficult task. The items should reflect the objectives as specified. Items should not be chosen just for the sake of generating questions. What is more, all items should be phrased clearly and should involve contextual information—they should not be testing language in isolation. And perhaps most important, tests should include several item-types. Research has shown that the test procedure, i.e. the type of item, may affect the performance of learners. Therefore, to neutralize this effect as much as possible, several testing procedures should be used. Another factor to be borne in mind is that items should be chosen which lend themselves to providing diagnostic feedback on returning the test. After all, a test should be a learning experience.

Finally, in considering the design of the test, we come back to the Genesis Syndrome. Even the most expert test designer must be ready to try out the test, and to modify or discard whatever does not work.

Testing Communicative Ability

1. Principles

Swain (1984: 17) sums up the communicative test in one sentence:

Each time you design or use a communicative theme-oriented curriculum unit, you have equally as well a potential testing unit.

Or, to put this in other words, the communicative test should be virtually indistinguishable from the communicative classroom activity. This will be seen in Swain's example that I will describe below.

Authenticity of language use is very important. Traditional tests have very little to do with real language in use, and this is reflected in some of the comments that Shohamy's learners made about their tests. The traditional test seems to be much more a stimulus-response routine produced in isolation than real day-to-day language. Communicative tests should present the language in context, and should be constructed in such a way that the test taker can use the context to his/her benefit, when doing the test—this is in keeping with Swain's (1984) principle of "bias for best." What is more, the communicative test, unlike its predecessors, should be integrative—it should not attempt, in the responses it is calling for, to tease out certain discrete features of the language. By not separating out discrete features, we are eliciting a much more real representation of the test-taker's language. However, it would only be fair to point out that

while this kind of integrative test meets the criteria for good communicative testing, scoring such a test and reporting and interpreting the results is a very difficult matter.

The communicative test should examine more than linguistic competence, both in the receptive and in the productive skills. The test should also be examining the learners' sociolinguistic competence—their knowledge of what Dell Hymes (1971) calls “the social rules of use”—the use of language appropriate to the medium, topic, setting, interpersonal relations and atmosphere. And, in addition, it should test what Canale and Swain (1980) call a “competence for discourse.” Second language learners have to be able to handle larger pieces of language and to recognize and be able to use the markers that hold a piece of discourse together, and signal what the logical relationships between propositions are.

I would like to describe an excellent communicative test developed for the French immersion programme by the Modern Language Centre at the Ontario Institute for Studies in Education. Swain's 1984 paper and Green's 1985 paper describe this test.

Before describing this excellent communicative test, Swain spells out her three underlying principles, and Green adds a fourth. They are:

- Start from somewhere
- Concentrate on content
- Bias for best
- Work for washback

Start from somewhere

By this, Swain suggests that the starting point in communicative teaching and testing should be a comprehensive model of communicative competence, and she advocates the Canale and Swain (1980) model of four competencies: grammatical, sociolinguistic, discourse and strategic competence. As she states: “having such a theoretical framework to start from is crucial” (p. 10).

Concentrate on content

First, the content must be sufficient to generate language that can be assessed for each of the four components of communicative competence. The content must be motivating in topic and presentation, the tasks must be real, and the students must learn something while doing the test. What is more, the content must be integrated and thematic and must foster interaction. As Swain (1984: 13) says:

In taking a communicatively-oriented test, the learner should have the experience of being communicated to, and of being able to communicate.

Bias for best

Swain's third principle is that in designing a communicative test, we should do everything possible to elicit the learner's best performance. This would include: encouraging the use of dictionaries and other reference material, allowing learners to work at their own speed with adequate time to finish; allowing the learners to review and modify work that they have already done; explicitly telling students what is being tested and looked for in each section; offering suggestions as to how to set about the task and clarifying who the audience is.

Work for washback

This principle calls on test designers always to bear in mind that the test will influence the teaching, and therefore to ensure that this influence is positive.

2. The O.I.S.E. French immersion example

These principles were put into effect in a test for French immersion students in grade 9. The test centres on a 12-page booklet entitled *A Vous la Parole*. The themes that were included were decided on after meeting with students of that age and finding out what would be relevant and interesting to them. The materials focussed on two summer employment opportunities—one was to work on a rock-concert series to be organized in the Francophone section of Sudbury, and the other was to tend vegetable gardens and farm animals in Fort Louisbourg in Nova Scotia.

Time does not permit me to give you a lot of the details, but let me mention some points that reflect what I have been discussing: teachers were involved with the project team at every stage of development; the student booklet is motivating, attractive and fun; the tasks are very realistic; writing a letter, a note for a bulletin board, a factual paragraph, an opinion composition, holding a conversation with peers and a job interview; there is what Swain calls "substantive content" on the two places, that the students learned from, and concentrated on.

Perhaps the best tribute I could pay to this project would be to say that *A Vous la Parole* would make an excellent *teaching* unit. Or, to use Swain's words (1984: 18):

When teaching and testing interlock in this way, the circle is no longer a vicious one—it is a positive and productive one.

3. Performance testing

Performance testing is a sub-set of communicative testing. More than any other communicative testing, performance testing has as its primary

objective to be able to *predict* how the test-taker will manage with similar tasks in the future under real conditions. Another important feature of performance testing is the emphasis that is placed on the test being as *real* as possible a simulation of the actual language tasks that await the test-taker. This attempt at approximating reality supercedes the objective of measuring general proficiency or of providing detailed diagnostic feedback to the test-taker (Wesche 1985). As Jones (1985: 16) states:

The identifying difference between applied performance and other types of tests is the degree to which testing procedures approximate the reality of the situation in which the actual task would be performed.

Performance tests are most appropriate for testing a specific target group for a specific purpose. An example of a situation that lends itself ideally to performance testing is university entrance language testing. There are a number of examples of such tests, such as the British Council's E.L.T.S. (English Language Testing Service) Test, and the O.T.E.S.L. (Ontario Test of English as a Second Language) Test, developed by a team headed by Mari Wesche, and of which I was a member.

A good performance test will be judged by the test-takers to be fair and representative, and the tasks will "look like" the real tasks they are attempting to simulate. For example, when working on the items for the Applied Science and Technology students in the O.T.E.S.L. Test, item types were decided upon more by looking at different first-year Engineering exams than by thinking of a variety of second language test items.

Performance testing should not be viewed as the domain of the experts, or as something that is only relevant to teachers of advanced level students.

There is an excellent example of a performance test of survival ESL described by Clark and Grognet (1985). This test, known as B.E.S.T. (Basic English Skills Test) was designed to assess the language of Indochinese refugees in the U.S.A. The test was developed specifically because, "the functional language requirements of refugees . . . constituted quite different instructional goals from those typical of more traditional ESL instruction." (Clark and Grognet 1985: 90).

4. Oral testing

This is the area of testing that teachers have been most hesitant about: it is extremely time-consuming, and, what is more, most teachers feel very inadequate in this area. The advent of communicative language teaching has compounded this sense of inadequacy—now there are sociolinguistic, discourse and strategic dimensions to be considered in addition to linguistic dimensions.

I would argue that the essential principle to be followed in classroom oral communicative testing is that the testing should be as direct and as natural as possible, and should not be very different from a good communicative activity in the talking class. The tasks should be contextualized and varied, and should generate language that enables the assessor to evaluate the sociolinguistic and discourse performance as well as the linguistic performance.

There is a common misconception that all oral evaluation will be both subjective and inaccurate. This does not have to be the case. If good rating scales are used (and there are several examples that can be followed such as Carroll's (1980), and OTESL's (1987)), and assessors are given training and practice, a relatively high degree of agreement can be achieved.

Another type of instrument that is extremely useful in assessing the sociolinguistic and discourse features of the learners' performance is the Communicative Checklist that we developed at the University of Toronto a few years ago for our Intensive ESL programme. This is not a rating scale, but rather a checklist of features to be thought about and listened for in the talking class. We use this as a teaching tool, placing it in the hands of both the teacher and the students, but it could easily be developed into a rating scale. Its main value is that it forces one to consider certain features of the spoken language such as paralinguistic appropriateness, sociolinguistic appropriateness, conversation management, etc., which are often neglected by classroom teachers who are used to listening for and addressing linguistic features only.

The oral test should be as rigorously planned as any other tests. All too often, oral tests do not test what the assessors need to know because of the careless, unstructured nature of the interview that is used. As a reaction to discrete-point traditional oral grammar tests, there was a move to unstructured interviews. Beginning with the blue-print of objectives, content and weight that I have advocated, one quickly realizes that other types of spoken language are needed in addition to the interview. What is more, the content of the interview itself requires careful thought.

There are numerous different types of oral tests, and, as I have said, more than one should be used. Role-plays and group discussions are very good for looking at sociolinguistic and discourse performance. Others include reporting, re-telling, describing a picture, chart etc., orally responding to a questionnaire, to name a few. Group discussions are very difficult to work with, but, from our experience, very effective in communicative oral classroom testing.

Times does not permit me to go into all the problems and difficulties of oral testing, but suffice it to say that I am acutely aware of the complexity and of such features as time, cost, reliability, fairness, authenticity,

etc. I would strongly urge teachers to read Underhill (1987) on this topic and to face oral testing with greater confidence in the future.

Concluding Comments

The central thesis of this paper has been that our testing should reflect our teaching—that the testing should be virtually indistinguishable from the teaching and should be a natural outgrowth from it. I would like to give the last word to Merrill Swain (1984: 13). This quotation encapsulates the essence of what I have been trying to say:

Communicative language testing and teaching are seen as two sides of the same coin . . . Having teaching and testing compatible is essential if we expect our students to learn what we teach them.

NOTES

1. This talk was presented as a plenary address at the TESL Canada/SPEAQ Conference in Quebec City, June 1988; and was accepted on the basis of blind review, as per the editorial policy of the *Journal*.
2. For the purposes of this paper, the term 'test' should be seen as not including the 'quiz'—the spot, unprepared-for classroom test or the ongoing informal testing that is part of every lesson.
3. The model I am proposing for test design is based on Shohamy (1985).

REFERENCES

- Canale, Michael and Merrill Swain.(1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics* 1. 1-47.
- Carroll, Brendan J. (1980). *Testing Communicative Performance*. Oxford: Pergamon Press.
- . (1985). Second language performance testing for university and professional contexts. In Hauptman (1985: 73-88).
- Clark, John L.D. and Allene Grognet (1985). Development and validation of a performance-based test of ESL 'survival skills'. In Hauptman (1985: 89-110).
- Cohen, Andrew D. (1980). *Testing Language Ability in the Classroom*. Rawley, Mass.: Newbury House.
- Green, Daina Z. (1985). Developing measures of communicative proficiency: A test for French immersion students in grades 9 and 10. In Hauptman (1985: 215-227).
- Hauptman, Philip, et al (Eds.). (1985). *Second Language Performance Testing*. Ottawa: University of Ottawa Press.
- Henning, Grant. (1987). *A Guide to Language Testing*. Cambridge, Mass.: Newbury House.

- Hymes, Dell. (1971). On communicative competence. In Pride and Holmes (1972: 269-293).
- Jones, Randall L. (1985). Second language performance testing: An overview. In Hauptman (1985: 15-24).
- Lado, Roberts. (1961). *Language Testing*. London: Longmans.
- Pride, J.B. and Janet Holmes (Eds.). (1972). *Sociolinguistics*. Harmondsworth: Penguin.
- Shohamy, Elana (1985). *A Practical Handbook in Language Testing for the Second Language Teacher*. Experimental Edition. Raanana, Israel: Shoshana Goldberg. (Soon to be published by Oxford University Press).
- Swain, Merrill. (1984). Teaching and testing communicatively. *TESL Talk* 15, 1 and 2. 7-18.
- Underhill, Nic. (1987). *Testing Spoken Language: A Handbook of Oral Testing Techniques*. Cambridge: Cambridge University Press.
- Wesche, Marjorie Bingham. (1981). Communicative testing in a second language. *Canadian Modern Language Review* 37 No. 3. 551-571.
- . (1985). Introduction. In Hauptman (1985: 1-12).
- Wesche, Marjorie Bingham, Michael Canale, Ellen Cray, Stan Jones, David Mendelsohn, Michael Tumpene, Marian Tyacke (1987). *The Ontario Test of English as a Second Language (OTESL)*. Toronto: Ontario Ministry of Colleges and Universities.

THE AUTHOR

David Mendelsohn is an Associate Professor of ESL at York University, and also an Instructor in the Woodsworth College TESL Certificate Programme at the University of Toronto. He is currently co-editor of the *TESL Canada Journal* with John Archibald. His interests include second language teaching and testing, with special emphasis on the aural/oral skills.