

LTJ_30_1_Assessing_Academic_English

[Start of recorded material]

Interviewer: From the University of Leicester in the United Kingdom, this is Glenn Fulcher with another issue of Language Testing Bites. In 2008 Cambridge University Press published *Assessing Academic English*, in which Alan Davies recounts the history of the international English Language testing system, known to language learners around the world as ILs. Along with the TOFEL, it is arguably one of the most widely used tests in the world. In issue 31 of the journal Christine Coombe reviews the book, and we thought this provided the ideal opportunity to talk to Professor Davies about the assessment of academic English.

Welcome to Language Testing Bites to talk about issues surrounding the assessment of academic English.

Respondent: Thank you very much, I am very glad to be asked to do this because it is an important issue and I think it needs more exposure perhaps than it has had recently.

Interviewer: In the review of your book, Christine Coombe agrees with the editors of the *Studies in Language Testing* series, that you were ideally suited to write this volume as you have been personally involved in the evolution of tests of academic English in the UK. Perhaps we could start by talking a little bit about the early days and the English proficiency test battery. Can you tell us why this test was needed then and why the assessment of academic English has grown into such a large industry today?

Respondent: It is kind of both Christine Coombes and of the editors to say what they have said. The fact is of course that I probably have a longer history in this area than many people, since I got started in it at the beginning of the 1960s, and so my experience goes back a long way. The position then in about 1960, 1961, was that the British Council were particularly concerned and involved with their programme in bringing overseas students, normally on British Council fellowships to this country, and they were concerned about their level of English so that they would then [unintelligible 00:02:27] properly at what they were doing here. The council itself had an existing test which was quite subjective and carried out by local officers in the host country where students were coming from. They were dissatisfied with this procedure because they felt that it was not at all reliable. They had different kinds of bias in different places. So the decision had been taken to develop an objective test of some kind, the University of Birmingham was involved with this because they had done some kind of report for the council, partly because Birmingham had a very well-known psychometrician, Edward Peel, and he was very much involved in that advice that they took on board. I happened to be in Birmingham at the time and in fact attached to the English department where I was working on a study of African writing in English. I had just come back after four years in Kenya, and my intention in fact, I had a small family at that time, my intention was to spend a period in Birmingham and then go back to Kenya, where we had been very happy indeed. I had been in some contact with the education department where Peel was, and had done some tutoring work for the Commonwealth [unintelligible 00:04:11]. So they knew about me, and they asked me if I would be interested in taking on this research project which was to develop an English proficiency test. We thought about it and decided that yes we would be.

So what happened was that I gave up the study of African writing in English and I started on this research project in conjunction with a very helpful official of the British Council, a man called George Heron, who had himself held a Simon fellowship in Manchester University, again he had also been working in East Africa looking at problems of overseas students in this country. He was very familiar with it and very supportive indeed I have to say. We set up the English teaching information centre at the British Council and Birmingham University gave me psychometric support, and so for the next two years I developed this test. It was very much a struck test, which was the vogue at that time. The council took it on, accepted it, it had quite good statistics, or satisfactory statistics anyway, and so for the next eighteen years they made use of this English

proficiency test battery which I had developed. Sometimes other people called it the Davies test, I never use that term myself. So that was the background to it, and it was a very Lado like test. I was influenced by Lado's book, as many of us were at the time, and so it was in some ways a strong American influence in the test.

Interviewer: As you discuss in the book, there came a point at which you knew that the EPTB content had been compromised, and that a new test was required in a response to the emergence of the communicative movement. The title of chapter two is the Communicative Interlude. What was it about the first English Language testing service that was communicative and why do you refer to it as an interlude?

Respondent: I'm afraid that in my previous response I missed out one point, but what I wanted to say was, because you had asked me about the development of the need for academic English. At the time I was involved in the early sixties, the number of overseas students here was very small, it was under 20,000 I think. By the early 2000s, it had gone up to over 300,000. So the growth was exponential. Now imagine it has gone up even more today, and that is true of course in other English speaking countries, Canada, New Zealand, Australia, and the United States. Now as far as the next question is concerned, this was about the communicative interlude. Such English language testing service, what was communicative about it you have asked. At the time this would have been in the mid to late 1970s, there was a strong influence of work on English, or indeed language, for specific posts, that was very fashionable at the time. The British Council in particular had become interested in that, partly because of the work of one of their officers, John Mumby who had written his PHD thesis which was precisely on that issue. They were influenced by that, one of his colleagues at the British Council, Brendan Carol, was particularly interested in it. And it was Brendan who took this one.

The communicative idea was, at the time, that what the test should do is to simulate, as far as possible, the demands placed on students when they were studying in this country in their subject areas. So there were two things really involved. One was the fact that they needed to interact with their teachers and so on, and the other more important one was that they needed to know the English of their subject area. English language testing service test was communicative largely because it set out to provide different tests for different academic areas. In point of fact, it didn't get further than six areas, and these six areas were related to the different departmental distinctions that the British Council itself when it was receiving students and assigning scholarships. So there was a test on medicine, for example. I know there was a test on engineering, and there was a test on social sciences. Why do I call it an interlude? Well the fact is that it proved to be very difficult to apply. The English language testing service was developed towards the end of the 1970s, and put into action and lasted for a further ten years. So it was actually put into use. It was, as I think I say in the book, in spite of its difficulties it demonstrated the kind of creative approach to testing, and also it showed the real difficulty of trying to test in exactly the same way as you might teach, required a larger amount of material in order to provide tests for all these different areas. I called this the problem of the logic of language specific testing. The problem was that it was necessary in a proficiency test, it divided into six or more, it was necessary to provide for all in those areas, and to do it there was a kind of regression to the mean as it were. In the sense that it was necessary to reduce difficulty, and the speciality of the area so that it fitted everybody in that area. A good example that is often quoted was that in medicine if you are going to have a test of medical English, what is it you base it on? You base it on texts for psychiatrists, for example, or for surgeons, or for physicians, or anaesthetists, or what is it you do? Because many of those who came here to study medicine [unintelligible 00:11:37] already gained qualifications in their own country, and indeed already begun to specialise in different areas. And some of them were years from general medicine, which is what they had studied in the first place, and so what they found was that the test was too specialist in that [unintelligible 00:12:01] and so the inevitable up was that in order to make a test apply to everybody taking it, it had to be very easy, rather like a Reader's Digest test. And that is what I call the problem, the

logic, or the illogic of the language specific tests, and also why it was an interview, because it was found during the 1980s that this was not satisfactory, and perhaps we will come to that later. Interviewer: Just to pick up on that particular issue. On page 39 you argue that the [unintelligible 00:12:34] experience did reveal that built in flaw in the logic of specific purposes testing. This echoes your argument in your 2001 article in Language Testing entitled the Logic of Testing Languages for Specific Purposes, in which you argue that language testing for specific purposes, and I quote, reduced to institutional definition of content instruction. Do you still believe this to be true, and do you think it is the primary reason for what you call the retreat from revolution in chapter 3?

Respondent: I think I do actually. I think that the logic of language specific testing is that you relate the test to an individual because individuals have different needs, and testing can't do that, not the kind of testing that we typically do anyway. Because the tests that we are involved with, especially proficiency tests, are aimed at groups rather than individuals.

Interviewer: That's very interesting, and I'd like to follow up by asking whether you think that language testing has helped us understand what academic language actually is?

Respondent: Now that is an interesting question. My response to that is that it has because in a way you design the test on the basis of your theoretical approach in what the domain consists of. And I think that is possible, and indeed it has been possible to do that for academic language testing. What you are saying, and indeed I'd like to quote here what I say in the book because I think this is crucial, what I say is that what is [unintelligible 00:14:38] can be generalised across subject disciplines, the generalisation involves argument, implication analysis, explanation, reporting, these I say are as true for literary studies as they are for accountancy, and for medicine and for all other academic disciplines. So, I think that was language testing has enabled us to do is to recognise that by using those, or by tapping into those skills and those abilities, we are getting at a [unintelligible 00:15:21] of requirement that [unintelligible 00:15:25] the academic spectrum. Of course it may be that there are other things that we are not getting to grips with, and that requires I think not so much further testing but further speculation, further analysis, further thinking as to what academic language really is. The irony of it all is that even when we have developed a test which involves some or all of those skills, we don't provide a prediction that is any better than the prediction we will get from a structural text.

Interviewer: Thank you. I'd like to finish the podcast now by raising probably the most important question around assessing academic English. The purpose of ILTs is to provide information on the readiness of the test taker for academic study in an English medium university. The score is therefore useful if, and only if, it has some predictive validity. Yet we know that ILTs, and indeed all tests of academic English have correlations with the criterion in the region of point three to point four. The amount of variants shared between test and criterion is roughly the same between ten and fifteen percent, irrespective of which test is used. So how can we say that the score is useful for its intended purpose?

Respondent: This is a question that is [unintelligible 00:16:59] isn't it. And I sympathise with it. Is it worth all the bother? My answer is this, that ten or fifteen percent, which as you say the variants between those figures is very generally found, and I accept that, I agree with that of course. That amount of variants, that amount of knowledge, that amount of prediction is worthwhile, and it is worthwhile because what is involved in a student's success, we take all number of variables. Variables that have to do with ability, with [unintelligible 00:17:38] with happiness, with finance, and so on. These are quite individual variables. If we can [unintelligible 00:17:51] one area which is right across the board, even though it is only ten to fifteen percent, that is worth doing. Point of fact, we don't want it to be more than ten to fifteen percent, because if it were that would mean that English ability in an academic area would benefit native speakers more than it should. It would mean that, let's assume it was 30%, well then if you were a native speaker you would be hugely advantaged because 30% of your academic ability would be down to the fact you are a native speaker of English. But of course we know that native speakers of English don't necessarily do any better, don't necessarily do well in their academic courses. It is

fortunate they know English of course, but it is the other variables that come into play, and particularly their academic ability. I'm quite happy with ten to fifteen percent.

Interviewer: Okay thank you. I've enjoyed reading your book as much as the reviewer has done. It is a fascinating history that raises a lot of questions that will vex the field for a long time to come. Thanks for agreeing to come on Language Testing Bites to discuss just a couple of those questions with us.

Respondent: Thank you very much.

Interviewer: Thank you for listening to this issue of Language Testing Bites. Languages Testing Bites is a production of the journal Language Testing from Sage publications. You can subscribe to Language Testing Bites through iTunes or you can download future issues from ltj.sagepub.com or from languagetesting.info. So until next time we hope you enjoy the current issue of Language Testing.

[End of recorded material]