

ENGLISH LANGUAGE TESTING IN ETHIOPIA: THE ESLC EXAMINATION

Harold S. Madsen

While Ethiopia's secondary school leaving examination serves as a useful yardstick in measuring the academic achievement of would-be graduates, it needs considerable refinement in order to achieve adequate reliability and efficiency. Perhaps no portion of the test is so woefully inadequate as the section dealing with English language skills; this despite the fact that examiners have considerably improved this instrument during the past decade.¹

Neither diagnostic nor truly comprehensive, the English language portion is highly subjective and generally ill-suited as an instrument for measuring the language proficiency of those whose native tongue is not English.

In fairness, it should be pointed out that one occasionally hears of worse examinations in both England and America: tests requiring parsing and intricate grammatical analysis, the regurgitation of rules and of tortuous terminology. The English language section of the E.S.L.C. avoids such idiocy. Furthermore, the reading passages and questions are culturally suitable for Ethiopians. Then too, the portions dealing with vocabulary and mastery of English idiom are well fashioned; disputed usage is avoided; incorrect alternatives have an authentic Amharic ring to them. Even the weakest portions of the examination reflect the good intentions of those who set the English section—a desire to see the language in *action* instead of a useless *analysis* of the language.

Unfortunately, good intentions are not enough. For instance, grim eighteenth and nineteenth century grammarians such as Robert Lowth, William Ward, Charles Coote, Lindley Murray, Thomas Dilworth, Caleb Williams, Samuel Kirkham, Gould Brown, Peter Bullions, William H. Wells and Samuel S. Greene strove in vain to produce symphonic eloquence through cacophonous rehearsals of syntactic-inflectional terminology.

First of all, the examination is not sufficiently comprehensive. It may be countered that the essay is intended to uncover deficiencies not tested in the objective section. However, given the freedom of an essay, students are able to slip into primer prose, if need be, to avoid being detected using faulty locutions. Aside from the essay and précis, this three-hour language examination contained in 1966 a scant fifty-four questions, thirty of which dealt with a few matters of idiom, the remainder dealing with such matters as vocabulary and written comprehension. An emphasis on mastery of idiom is sound; however, to insure validity and reliability, those who set the examination should provide several times this many questions. Also, if a comprehensive measure of language proficiency is desired, then many areas of language usage need to be tested.

Besides its lack of comprehensiveness, and partly *because* of this deficiency, the examination is not truly diagnostic. One can tell that the patient has a fever, but whether the malady is measles or malaria is somewhat conjectural. Numerous, carefully selected questions covering a wide range of language matters would provide a

1. And no one is more anxious to see significant additional improvement in the test than the present Chief Examiner in English.

better opportunity for diagnosis and consequently more intelligent solutions for prevention and cure.

A third weakness of the examination is its appalling subjectivity. Barely over 25 per cent of the 1966 test was immune from examiner bias. Over 50 per cent of the student's score on the English language portion of the E.S.L.C. was determined by the subjective evaluation of the essay and précis. The inevitably low correlation between examiner grades not only casts serious doubt on the evaluation of précis and essay but also tends to invalidate the entire English language test, except for scores at extreme ends of the scale.

Finally, the composition and précis (constituting 55 per cent of the English test) are not suitable instruments to measure the language ability of those for whom English is a second or even third language. In brief, experts in the field of teaching English as a second language deplore the assignment of long, free compositions. And modern linguists as well as test experts challenge the efficacy of précis writing as a measure of language proficiency. The essay and the précis are undoubtedly the least defensible items on the entire three-hour English language test.

A long uncontrolled composition, as required on the E.S.L.C. exam, presents several serious difficulties.

TESL specialists hold that controlled, semi-structured, or very short writing activities are more suitable for those learning English as a second language than are long free compositions. While long compositions are not felt suitable for freshmen English students at Haile Sellassie I University, a long composition is nevertheless required of E.S.L.C. candidates. The results, of course, are disastrous. Students simply make the same mistakes a dozen times over. Moreover, they wander aimlessly like a delirious person in strange woods at midnight. In the recently published report of the 1966 E.S.L.C.E., English essay examiners found "total lack of planning or any form of coherence." These examiners asserted that even though "the standard required for a pass in the essay is not an impossibly high one ... *the great majority of the candidates fall so far below the required standard that their attempt at this section of the examination is a sheer waste of time both for them and the examiners [my italics].*" (page 52)

Another serious problem with the essay is the extreme subjectivity of examiner grading. Four years ago in Great Britain E.L. Black published a study of the marking of GCE scripts. Nineteen paid examiners from one examining board were given special marking instructions at a co-ordination meeting. Then they were handed identical essays to be graded on an eighty-point scale. Scores ranged from 51 per cent to 34 per cent. When the experiment was repeated the following year, percentages ranged from 75 to 38; the third year, from 61 to 44; and the fourth year from 62 per cent to 39 per cent. Commenting on these and other findings, another recognized British test specialist says, "With the best will in the world, one finds it difficult to place any faith in the marking of compositions in the face of these figures."²

In America, too, the same findings have been made again and again. As far back as the 1880's this otherwise ideal testing instrument was discovered to be fraught with "the pitfall of unreliability." And over the years mountains of evidence

2. D. W. Grieve, "English Language Examining," Lagos, Nigeria: African Universities Press, Ltd., 1964; pp. 40-41.

have accumulated. Typical was a study at the University of Tennessee in 1951. Worried about the efficacy of test essays, the English department compared 1,021 such essays (as well as student scores on a national standardized examination) with course grades in English at the end of the quarter. The essay was so unreliable in locating superior and poor students that it was moved to the first class period, and the objective test became the main criterion for placement.³ In 1956 even the College Entrance Examination Board decided to discontinue their attempts to provide a reliable marking of test essays for university admissions.

A great deal more needs to be said about the difficulties of the long, uncontrolled essay as a test of language proficiency, but a mere summary will have to suffice for the present. The essay test in its present form on the E.S.L.C.E. is inappropriate for those whose native tongue is not English. Secondly, such a test often lacks validity since a single writing stint may fail to reveal one's actual capacity to express himself. Any teacher of composition recognizes how widely student performance on essays varies, due to the nature of the subject and the frame of mind of the student. Finally, the essay is pathetically unreliable. This follows in part because dozens of skills are being tested simultaneously: skills ranging from mechanics, grammatical concord, sentence structure and idiom to diction, organization, style and logic. True, the examiner may attempt to restrict his analysis, but even so the weighting of errors is extremely difficult and extraneous matters inevitably thrust themselves upon him. Moreover, the essay is unreliable as a test instrument because individual examiners are not entirely consistent in their grading from week to week, and the correlation of examiner marks is low. As indicated earlier this lack of consistency tends to invalidate not only the essay portion but also the entire English language section, since the essay constitutes 35 per cent of the English grade.

The précis is a highly technical accomplishment which has no business being taught to learners of English as a second language—at least not below the college level, and even here it has very, very little to recommend it. Presently it seems to be avoided by all English teachers at the University. As a measure of general language ability it is virtually useless.

Nevertheless, it is more impervious to the findings of linguistic science than blood-letting, the wearing of charms, astrology, and alchemy have been to research in medicine and the physical sciences. Précis writing has enjoyed a formal tradition in much of the English-speaking world since the reign of Queen Anne (1700-1712) when the undersecretaries of diplomats were required to condense long documents in order to simplify business proceedings. Currently it is advocated in order to enable writers to distill vast amounts of material (a rare need), to discipline the mind (a psychological concept rejected in the nineteenth century), to promote accurate reading, to develop concise and direct writing, and to assist students to discriminate between principal and subordinate ideas. The latter are legitimate objectives, but précis writing is an artificial and ineffective means of achieving these ends.

The now famous "Grieve Report" (a very searching, scholarly, objective evaluation of the West-African equivalent of the E.S.L.C.) completely repudiates the précis test as a valid examination of English language skills. Following an extensive study of this examination, D. W. Grieve in 1964 published his findings; the following eight charges are taken from his report: The précis test "is very imperfect, because it is testing two skills at the same time: comprehension and continuous writing."

3. K. L. Knickerbocker, "Placement of Freshmen in First Quarter English," *Journal of Higher Education*, XXII (April, 1951), 211-215.

Second, as a test instrument it is a "poor discriminator." In addition it has a bad effect on teaching—too much time being spent on "techniques" and not enough on learning to read and write English. It encourages the teaching of various "tricks of the trade" and "highly artificial forms of language use." His fifth allegation is that the arbitrary requirement of a one-third reduction is unrealistic. Equally so, said Grieve, is insistence upon a paraphrase. Seventh, because of précis difficulty, teachers spend a great amount of time on it; but "for all *practical* purposes, this is time wasted." Finally, "like composition, it is unreliable." He ends his comments on the précis test by recommending that it "be eliminated as soon as may be convenient."⁴

That such a step should also be taken in Ethiopia is hinted at by the English examiners of the 1966 E.S.L.C.E. In their report released the last of October, 1966, they say:

It would appear, however, that the time is ripe for a reappraisal of this particular section of the English Language Exam. What exactly is the purpose and the value of précis-writing are questions about which there is some disagreement among English teachers. Where there is no uniformity of opinion among teachers, how can we use this test as a fair means of evaluating candidates? Précis-writing appears to have originated in examinations for candidates for the English Civil Service. These candidates were required to show their ability to read a passage (sometimes of several pages in length) and reproduce in précis form its main points. This traditional examination, dating from at least 50 years ago has been perpetrated [*sic*] all over the world in English examinations, based on the British system, and it shows no sign of dying off of old age. The value of the précis is that it requires a candidate to show his understanding of a passage and his discernment of the relative importance of the different points made in it. Of course, if a student cannot write even one English sentence correctly, he can scarcely reveal his understanding of a passage.

We have now considered some deficiencies not only of the précis examination but also of other portions of the E.S.L.C. test in the English language. Obviously, immediate steps need to be taken to improve the examination.

A committee of experienced English teachers, in cooperation with persons qualified in testing and appropriate ministry people, need to re-evaluate and re-write the exam.

A good revision would undoubtedly produce an instrument that is far more objective. It is to be hoped that this can be accomplished without undue hysteria or name-calling. Simply to dismiss objective tests as "American" or to charge that "you can't teach students to write by just circling correct answers" is irrational. The latter reflects a confusion of diagnosis and remedy. Suitable writing assignments in the classroom *are* indispensable. But Paul B. Diederich, renowned expert in essay and composition testing in English, counseled English teachers as early as 1949 that there is "more sense" in using objective tests than most realize, "because objective tests on reading and writing can now be made sufficiently subtle and penetrating to disclose the higher levels of skill in these arts."⁵ A revised E.S.L.C. English test would likely include far more questions and cover far more areas of English than at present. If the essay were retained, it would likely be much abbreviated and rather strictly controlled, as suggested by Grieve. The précis test would simply be eliminated from all subsequent examinations.

The youth of Ethiopia deserve not only better *training* in English (a goal yet a long ways off) but also more careful and honest *appraisal*—a goal easily within grasp.

4. Grieve. *op. cit.*, pp. 47-48.

5. Paul B. Diederich, "The Use of Essays to Measure Improvement," *College English*, X (April, 1949), 396.