

---

---

## NEW DIRECTIONS IN THE TESTING OF ESL WRITERS

**Speakers:** Ulla Connor, Indiana University  
Joy Reid, Colorado State University

**Introducer/Recorder:** Dennis Donahue, New Jersey  
Institute of Technology

Joy Reid and Ulla Connor explained two new analytic systems that describe and evaluate ESL writing. Reid presented a system that can be used to examine essays on the basis of syntactic complexity, and Connor described a method of understanding the components of persuasive essays. The presenters offered their systems as methods of evaluating ESL students' written work and of improving composition instruction.

To ascertain whether quantitative differences exist in essays written by different groups of ESL students, Reid used a computer text-analysis program, the *Writer's Workbench*, developed by AT&T Bell Laboratories. She examined more than 750 essays in her study, measuring such areas as word length, percentage of long and short sentences, use of passive voice, readability, and percentage of pronouns. Two different types of writing topics were used--comparison-contrast and interpretation of a graph--and there were two questions within each essay type. In addition, the essays were examined for differences related to the students' language background (Arabic, Spanish, Chinese, and English), gender, and major field of study, and for differences resulting from the three different holistic scoring systems used to score the essays.

Reid discussed some of her findings, the most significant of which was that different language groups frequently write differently. For example, Arabic-speaking students wrote the longest sentences (though comma splices may have contributed to this finding), and native speakers of English used more pronouns and used the passive voice more often than students in the other three groups. Reid also discovered that there were quantitative differences between essays written on different types of topics. The essays explaining the graphs, she found, were often longer than the comparison-contrast essays. Reid distributed a multi-page

booklet presenting her results in tabular form to those present.

Next, Connor reported on her attempt to find a method that would give more specific indications of writing progress than holistic assessment alone does. She noted that students may make progress during a term of instruction but they often receive the same holistic score that they received at an earlier testing. To be able to document students' progress in persuasive writing, Connor used several methods of analysis. She mentioned two in summary form--her examination of syntactic features using the computerized analysis of 39 syntactic variables, developed by D. Biber at the University of Southern California; and her study of coherence, using a scale that includes nine different variables related to topic development. Her emphasis, however, was on a third method, one that can be used to evaluate an ESL writer's control of informal reasoning.

Connor explained her use of Stephen Toulmin's Criteria for the appraisal of informal reasoning in written work. She has used this system of analysis to evaluate strengths and weaknesses in a student's writing that cannot be assessed by holistic scoring. This system makes it possible to explain the differences in writing skills of two students who receive the same score on a writing sample rated holistically. Connor and a colleague examined the responses of twenty-six students to the question asked on the Fall 1986 TOEFL Test of Written English (TWE). They were looking in particular for elements Toulmin described as claims, data, and warrants. In the rating scale used, each of these elements is given a score ranging from 1 (lowest) to 3 (highest). Claims are judged according to how specifically they are stated; by the number of relevant subclaims developed; and by how logically, feasibly, and originally the claims are developed in the essay. Data are rated on their quantity and quality. Warrants--judgments made by writers that are based on their claims and data--are rated according to their number, their reliability, and their relevance to the claim made. Connor stated that test-takers in general received higher scores for claim and data than for warrants, which require higher level thinking skills.

Connor noted that the methods of analysis that she and her colleague used allowed them to measure changes in writing ability that holistic scoring alone cannot measure. She distributed to those attending the session a copy of a student response to the TWE topic and a Toulmin analysis of that response.

The speakers concluded by noting that the assessment methods that they had presented could be used for ESL writers and for native speakers.◊