PROCEEDINGS
VOL. II  PAPERS ON APPLIED LINGUISTICS

THESALONIKI GREECE
APRIL 2-4 1998

DEPARTMENT OF THEORETICAL AND APPLIED LINGUISTIC
SCHOOL OF ENGLISH LANGUAGE AND LITERATURE
FACULTY OF PHILOSOPHY
ARISTOTLE UNIVERSITY OF THESALONIKI

OR: S. LAMBROPOLOU
Testing and Comparing the Writing and Speaking Ability of EFL Students.

P. PAVLOU, N. CHRISTODOULOU, P. ZARPETEA
Intercollege, Nicosia, Cyprus

ABSTRACT

This paper examines two predominant theories of language competence and their implications in terms of teaching and testing. The aim of the study is to evaluate and compare a learner’s writing and speaking ability in order to establish whether these two distinct abilities represent a single unitary ability or two different abilities. The results of this study could enable language teachers form homogeneous groups according to specific language skills.

THEORETICAL BACKGROUND

In language studies it is fundamentally essential to establish what it means to know a language. The answer to this question can serve as a guide to teachers, applied linguists, language testers etc. in deciding what to teach or test.

There are many schools of thought that have attempted to define language or language competence. One of the two predominant theories of language competence is the one represented mainly by John Oller and supports that language competence is a unified set of interacting abilities which cannot be separated and tested adequately. This theory is also known as unitary trait hypothesis. According to this theory, language competence consists of a single unitary ability. Using a variety of language tests and correlating their scores Oller claimed that he discovered a “g-factor” which he regarded as a unitary trait called “general language proficiency”. However, later studies showed that the unitary trait hypothesis, in its strongest form, was wrong (Oller 1983).
The other main school of thought represented by, among others, Canale and Swain (1980) and Bachman (1990) revolves around the assumption that language can be broken down into smaller components. These components are the four skills (listening, speaking, reading, writing), the various hierarchical units of language (phonology, morphology, syntax etc.) within each skill and subcategories within those units i.e. affixes, tenses, pronunciation, intonation, sentence structure etc.

TEACHING IMPLICATIONS

What are the implications of these two theories in terms of teaching? Teachers supporting the first theory would adopt a more integrated approach in their teaching methods and avoid teaching the four language skills in isolation.

On the other hand, proponents of the second theory would adopt an approach focusing on the various skills separately. Since they recognize that language competency consists of various distinct components, they consciously aim at developing the various language skills.

LANGUAGE COMPETENCY AND TESTING

If the first theory is adopted, then a learner’s performance on one skill or on one linguistic unit (syntax, vocabulary, grammar) would be considered as adequate indication of the learner’s competence on all four skills or on all linguistic units.

For example, a test on writing could be used as a tool for making judgments on the learner’s overall ability in a foreign language including listening, speaking and reading.

Such assumptions, however, are not allowed by the second theory. A test of speaking based on this theory would provide us with information only on the learner’s competency to use spoken language and not on the other skills.
THE STUDY

Based on the above drawn dichotomies we wanted to examine the validity of the current placement policy of our institution (Intercollage). Intercollage is an institution of tertiary education where the language of instruction is English. According to this policy, new students are tested by means of a written test. The results on this test are supposed to indicate the students’ overall ability in English.

Students who perform similarly on this test and therefore are assumed to have the same overall ability are placed in the same level. However, we have been faced with the problem of having multi-level classes, that is, there appears to be no homogeneity in terms of the students’ ability to communicate in English. Therefore, in a multi-level class special problems arise that the placement test was supposed to prevent. Such problems are the inability to follow the course outline and the necessity to repeatedly teach certain skills, thus making the learning experience less interesting and less challenging for some students.

These problems made us suspect that there could be certain limitations in the placement process. We suspected that part of the problem was that the placement test consisted of only a written component. The absence of an oral component in the placement test was thought to contribute to the assumed misplacement of certain students and hence the creation of multi-level classes.

We decided to test our assumption by means of a research project.

METHODOLOGY

In order to carry out this research project the following methodology was adopted.

The written and oral samples of 37 EFL students from various linguistic backgrounds at Intercollage were used. The written sample was the Placement Test which has been used at Intercollage to assess the students’ overall proficiency in order to place them in
one of the six courses (currently offered by the English Department) which represent various levels of proficiency in English. These courses are:

Beng-050: For beginners with no or little knowledge of English. (15 hours a week)
Beng-080: Roughly equivalent to lower intermediate (12 hours)
Beng-090: Equivalent to intermediate (8 hours)
Beng-100: Upper Intermediate (6 hours)
Eng - 100: Advanced (3 hours)
Eng - 101: Superior (3 hours)

The first 3 courses are very intensive and are basically grammar courses for EFL students.

Beng - 100 is a transitional course which combines basic sentence structures and advanced grammar, as well as paragraph and composition organization and reading skills (such as skimming and scanning).

The last two courses on the scale (Eng-100, Eng -101) are comparable to courses offered to students (native and non-native) at American Universities and are not considered EFL courses.

The placement test consists of two parts. The first one is a multiple choice test focusing on grammar, vocabulary and reading comprehension, and contains a total of 70 questions modeled after the TOEFL. The second part is a composition. The student can write on one of three topics each representing a different type of writing (narrative, descriptive, persuasive, argumentative, compare contrast and process analysis). Emphasis is placed on narration, description and argumentation. Students are allowed one hour for the objective part of the exam and half an hour for the composition. This composition is evaluated on a 1-5 scale. This is the scale that Intercollege has adopted and used. The descriptors of this scale are very general. Also, there is a breakdown of the multiple choice scores according to level.

The second tool used in our research was an oral interview. The oral samples of the students were elicited by an Oral Proficiency test in the form of an interview. There were two sets of questions.
One was geared towards novice students and the other towards more advanced students. Very early in the interview the examiner determined which set to use based on the initial impression of the interviewee.

There were two sets of questions of the oral interviews reflecting the structures that are tested in the objective part of the written test (mainly grammar). This was done because our aim was to test exactly the same structures through two different mediums namely written and spoken. The difference between the two sets of questions was that the first set consisted of yes-no questions whereas the second set was composed of open-ended questions. Some of the questions in the second set aimed at eliciting extended speech by higher level students.

**SCALE**

Both the written and oral samples were assessed against a rating scale developed by the researchers. This rating scale has descriptors of grammar, vocabulary and syntax distributed over six levels ranging from novice to superior. This rating scale is more detailed and analytic. Throughout the study, this scale will be referred to as the new scale.

The structures appearing at each level are those that a student is expected to master in order to move to the next level. This also means that this student has already mastered the skills of the previous level (levels).

**SAMPLE (THE SUBJECTS)**

Five to eight students from each level were randomly selected. The sample consisted of 37 students. Each subject was interviewed for 15-20 minutes. All the interviews were videotaped.

The grades given to the oral and written samples were entered into the EXCEL computer program which is a data storage software and automatically runs basic statistical operations such as
correlations. More specifically, the grade of the oral interview was correlated with the grade given to the composition (new scale) and the grade on the multiple choice section of the placement test. In addition, the grade on the oral Interview was correlated with the overall score on the placement test for each student. Finally, the grade on the composition at the placement test was compared to a new grade on the same composition. The new grade was obtained by using the scale used to evaluate the oral interview samples.

The interviews were conducted by three interviewers and were subsequently rated by them.

RESULTS

Comparison of Oral Interview Vs Comp. (New Scale) (BC)

When comparing the score on the oral interview to the score on the composition using the new scale the correlation is high (0.83). One of the reasons that explains this high correlation is the scale itself. This new scale contains descriptors that are more detailed and analytic and reflects the same structures. This result supports Oller's unitary trait hypothesis namely that language ability is a unified whole and it is manifested equally in the separate skills. In our study, it is evident that the students' speaking ability is compatible with their written ability.

Comparison of composition scores using 2 different scales (CD)

When comparing the composition scores obtained by the two different scales the correlation is again very high (0.93). Even though we had suspected that the holistic nature of the existing scale did not differentiate students' writing ability well, the new analytic scale yielded very similar results. The reason we created this new scale is because we wanted to be able to evaluate the compositions more objectively hoping to achieve a greater differentiation among
the students. However, the results seem to remain the same regardless of whether the compositions were graded holistically (previous scale) or objectively (new scale).

Comparing the Composition Score on Placement Test (previous scale) Vs multiple choice score on Placement Test (DE)

When comparing the composition score on the placement test (previous scale) the correlation is very high (0.91). Basically, this tells us that if we classify students according to either their score on the composition or the multiple-choice, the student will be placed at the same level. Therefore, strictly speaking one of the two components of the placement test becomes obsolete.

Comparing the score on Oral Interview with students’ overall score on placement test (BF)

When comparing the score on the oral interview with the students’ overall score on the placement test the correlation is very high (0.9). Once again this proves that the performance on a specific skill (i.e. spoken ability) is a good/accurate indicator of the student’s overall ability. Therefore, if we use only the score on the oral interview we can place the student fairly accurately.

CONCLUSION

The major conclusion of the results of this study is that the learner’s overall language competence is a unified whole and exemplifies itself similarly in the various skills.

SUGGESTIONS/RECOMMENDATIONS

Even though the three different ways of testing a student’s overall language ability (i.e. oral interview, composition, multiple-
choice) are equally reliable and sufficient on their own, each one is useful for the following reasons:

A) Information derived by these three different ways of testing will be helpful in creating courses of the same level that would have different focus. For example, the same advanced level course could have two sections focusing on the improvement of specific aspects of a skill. Thus, a student would have the opportunity to improve his/her pronunciation (speaking) or his/her ability to create a coherent text (writing) while the core of the course remains the same. By the same token, reading techniques (i.e. skimming and scanning) could be reinforced for students who are weak in these areas.

B) The use of a variety of testing tools gives the opportunity to accommodate students’ different personalities. For example, a student may perform very well on the writing component of the test but may be too timid to participate in the oral interview. Or, someone may not be comfortable to respond to multiple-choice questions.

IMPLICATIONS OF THE STUDY ON THE PLACEMENT PROCESS

In our introduction we stated that one of the aims of our study was to examine whether the existing placement test was one of the causes for the creation of multilevel classes. We had hoped that by instituting an oral component in the placement process would assist us in obtaining a more accurate picture of a student’s language competence and consequently create more homogeneous classes. However, the results show that despite the use of three different testing tools, the problem of having multilevel classes would not be eliminated.

As English teachers, we feel that further examination of the matter is needed. We want to consider other possible causes of the problem.

It is evident from our study that the students do not differ
substantially in their oral and written performance as measured against the new rating scale. Therefore, the reasons for the differences in language ability among the students must be looked elsewhere.

For example, throughout our interviews we came across students who showed mastery of grammar structures but were virtually incomprehensible. Thus, if we used only the oral component as a placement tool, the student would most likely be placed at a lower level class where he/she would be familiar with the material taught.

By the same token, if we used only the written component, the student could be placed in a class with students who are very fluent in spoken English, in which case his English would not be compatible with the rest of the class.

Therefore, we feel that a fluency component should be added to the evaluation/placement process.

Another reason for the existence of multilevel classes is the absence of rating criteria relevant to the textual organization of compositions (i.e. coherence, unity, paragraphing). This becomes more crucial in the placement of students in higher level classes (Eng - 100 and Eng - 101). However, in the case of low level students the writing is so limited that these criteria cannot be applied. Therefore, unavoidably more emphasis is placed on the objective part of the test.

For example, a student may be in very good command of the grammar and be placed in a high level course where he would be expected to be able to produce a coherent and well - organized text. Since the placement test does not evaluate such ability, the student would seem weak compared to other students.

One way to remedy this situation is to evaluate the composition in two ways: a) to score it analytically focusing on the grammar structures and b) to score it holistically focusing on the rhetorical organization, content and communicative expression.

Finally, another recommendation which pertains to the composition part is the inclusion of topics that would allow the student to show mastery of a variety of structures. For example, a student may choose to write a narrative using the past tense without having to show that he is in command of other structures. A
possible remedy would be to ask the student to write three shorter writing pieces each one reflecting different structures.

### APPENDIX

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>15</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>14</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>22</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>21</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>25</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>27</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>21</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>24</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>20</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>48</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>17</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>26</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>38</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>19</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>41</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>27</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>21</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>40</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>22</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>43</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>23</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>34</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>24</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>36</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>25</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>53</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>26</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>50</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>27</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>45</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>28</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>38</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

A=Student  
B=Score at Oral Interview  
C=Composition Score with New Scale  
D=Composition Score on Placement Test  
E=Score on Multiple Choice  
F=Overall Score at Placement Test


<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>55</td>
<td>3</td>
</tr>
<tr>
<td>30</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>50</td>
<td>3</td>
</tr>
<tr>
<td>31</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>47</td>
<td>3</td>
</tr>
<tr>
<td>32</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>55</td>
<td>4</td>
</tr>
<tr>
<td>33</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>60</td>
<td>5</td>
</tr>
<tr>
<td>34</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>52</td>
<td>4</td>
</tr>
<tr>
<td>35</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>62</td>
<td>5</td>
</tr>
<tr>
<td>36</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>64</td>
<td>4</td>
</tr>
<tr>
<td>37</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>52</td>
<td>3</td>
</tr>
</tbody>
</table>

REFERENCES