Assessment of the EFL Learners' Dictionary Using Skills

Yukio TONO
(Tokyo Metropolitan College of Aeronautical Engineering)
Assessment of the EFL Learners' Dictionary Using Skills

Yukio Tono
Tokyo Metropolitan College of Aeronautical Engineering

1. INTRODUCTION

Lexicography is a field concerned with the theory and practice of dictionary-making. It has a long tradition and the number of its products is extremely large. Recently, however, this field of lexicography has been changed in its perspectives. Lexicographers used to make dictionaries according to their own ideas of what should be in a dictionary and the users received whatever they were given. It is true that these dictionaries were reviewed in the journals and the newspapers everytime they were published, but no systematic research on how the real users used them and what kind of information they needed more or failed to retrieve had been done until a decade ago. There has been a growing awareness that there is a gap between the lexicographer's own expectation of what a dictionary should be and the user's real reference needs and skills. The user perspective in the study of lexicography is a relatively new field, but it has become one of the most flourishing areas in lexicography theory, research, and practice.

In this paper, we will focus ourselves on one of the most important but not very much investigated aspects of the user study, that is, the assessment of dictionary using skills. First related literature will be reviewed in Chapter 2. Then the research
design and its results and discussions will follow in later chapters.

2. TWO MAJOR RESEARCHES ON TESTING DICTIONARY USING SKILLS

There are very few systematic researches on testing dictionary using skills. As is illustrated in Hartmann (1987a), most of the user study is concerned about the following four aspects: 1) dictionary typology (investigating what should be in a dictionary), 2) user typology (studying who needs dictionaries), 3) needs typology (analysing the user's needs), and 4) skills typology (studying how the user finds what he/she needs).

Our topic might be related to the fourth category most closely, but most of the studies in this category are concerned about HOW the user finds what he/she needs and not HOW WELL he/she does so. Since the user study has just started recently, we know very little about the cognitive strategies of dictionary look-up. It is true that assessing the dictionary using skills is difficult when we know very little about what takes place during the user's dictionary look-up, but it is also possible and important to develop such an assessment test with which we could investigate the dictionary using skills and have better understanding of the whole process of dictionary look-up.

In this section, we are going to review the methods of the two major studies about testing the dictionary using skills, one in Great Britain and another in Japan, and consider some major factors in assessing the dictionary using skills.
2.1. Atkins et al. (1987)

The first one is the research project undertaken by Beryl T. Atkins and her colleagues (1987). This project is very ambitious because what they try to do is to know every aspect of the use of learner's dictionaries through the empirical data obtained from all over the world. The purpose of their research project is "to discover how effective a learner the student of English as a foreign learner[sic] is when working with a bilingual and/or monolingual dictionary." (1987:29) This is still the on-going research project and the findings will be reported eventually.

In this preliminary report, they set the three-fold research tools. First, they proposed to ask students to complete a User Profile Questionnaire in their own language. This was to be done in order to get the information of the user, which might not be obtained from the test (for instance, any experience of receiving instructions in the use of dictionaries and so on). Second, Classification Tests would be given in order to classify the subjects on the basis of their English language skills. Then, third, Assessment Tests would be given to assess their dictionary-using skills and to evaluate the selection and presentation of material in the dictionaries used (ibid.).

The third part of their research design is directly relevant to our present topic and the contents will be reviewed in a later chapter. Meanwhile, we can go through their research aims. They claimed that the research aims were to "discover how much instruction students receive in the use of dic-
tionaries, and is also designed to provide a comparative evaluation of the effectiveness of bilingual dictionaries and monolingual learners' dictionaries for carrying out various operations in language use and language learning." (Ibid.) Another purpose is "to find out how dictionaries fail students, and hence how dictionaries may be improved..." (Ibid.)

Here we can clearly see the idea that dictionaries can be and should be improved with the data obtained from the users directly. Of course the users have to be educated so as to be able to get the most out of dictionaries, but at the same time dictionaries should be improved in terms of user-friendliness and easy access to the data.

Atkins and her colleagues identified the following nine factors to be involved in testing the dictionary using skills (p. 33):

(1) DICTIONARY TYPE: whether it is monolingual or bilingual.

(2) ACTUAL DICTIONARY USED: the specific dictionary being used and hence the actual content and organization of the entry that the student is consulting.

(3) OPERATOR (A): first factor relating to the user, viz. his/her intelligence, linguistic ability and knowledge of English.

(4) OPERATOR (B): second factor, viz. the user's general dictionary-using skills.

(5) OPERATOR (C): third factor, viz. the user's familiarity with the particular dictionary being used.

(6) OPERATION: the actual operation that the dictionary is being used for, viz. decoding and en-
coding tasks of various types and language learning activities.

(7) PRESENTATION: the way in which each dictionary variously presents the information.

(8) MATERIAL: the actual type of language material required for the task in hand.

(9) TEST: the format and type of the specific test being used.

Each of these factors should be treated in the following way:

Treated in the assessment test: ... (1), (4), (6), (7)
Treated in the questionnaire: ... (2), (5)
Treated in the classification test: ... (3)
Controlled by the actual test being undertaken: ... (8), (9)

Therefore, the following identification and control of the types of test to be devised is possible:

<table>
<thead>
<tr>
<th>Items to be varied for testing</th>
<th>Factors which must be identical, or as similar as possible</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2/3/4/5/6/7/8/9</td>
</tr>
<tr>
<td>4</td>
<td>1/2/3/4/5/6/7/8/9</td>
</tr>
<tr>
<td>6</td>
<td>1/2/3/4/5/7/8/9</td>
</tr>
<tr>
<td>7</td>
<td>1/2/3/4/5/6/8/9</td>
</tr>
<tr>
<td>8</td>
<td>1/2/3/4/5/6/7/9</td>
</tr>
</tbody>
</table>

We can now think of some patterns of tests as follows:

1 Dict type          [+MONOLINGUAL]/[-MONOLINGUAL]
4 Operator           [+ADVANCED]/[+INTERMEDIATE]/
The combination of these variables will give us an exhaustive list of the tests of dictionary use. The tests are designed to allow cross-comparison of results. For instance, all other variables being constant, the pair of the tests with [+MONOLINGUAL] and [-MONOLINGUAL] will give a comparison of the effectiveness of monolingual and bilingual dictionaries in the same context.

The research design is very solid and the eventual findings will be of considerable importance to lexicographers and teachers. Methodologically, however, feasibility is the most biggest problem of this project. It will be a tremendous work to make every set of assessment tests and this will not be possible without help of many different organizations or individuals undertaking the development of the test. It would be very difficult to ask different persons from different countries to make a test and to compare one test with another, assuming that all the variables but one must be constant.

2.2. Okayama (1985)

This was a part of the series of studies of teaching English in high schools in Okayama, undertaken by Okayama Prefecture Senior High School Circle of English Education. This study was unique in the sense that they developed what they called "Dictionary Using Skills Test" and administered it to the high
school students in Okayama. As far as the present writer knows, this is the only report of the results of assessing the real users' dictionary using skills.

The purpose of the study was to see how effectively the learners of English could use English-Japanese dictionaries. Even though in the questionnaires the students answered yes to the question of whether they could use dictionaries well, most of the high school teachers felt that the students had difficulty finding words and appropriate meanings in dictionaries. Thus they decided to develop a test which could evaluate the students' dictionary skills objectively.

The scope of the battery was very narrow, as compared with the project by Atkins, et al. (1987), since they made a test assessing only the skills of using bilingual dictionaries for a recognition purpose. However, this study is important because it actually gave us the results of the tests, which will lend insight into the development of more valid and reliable tests in the future.

The contents of the test will be compared with the one by Atkins et al. (1987) in the next chapter. The results of the tests will also be discussed in later chapters.

3. THE COMPONENTS OF THE DICTIONARY USING SKILLS TEST

Let us now consider what should be included in the assessment test of dictionary using skills. Atkins et al. (1987) pointed out that the following aspects of dictionary use should be included as the components of the test (p. 37):

(a) points of detail which are the same for both
monolingual and bilingual dictionaries, e.g. understanding of grammar and grammar metalanguage; where user expects to find compounds, idiomatic phrases, etc.; morphological variations; subsidiary information about headword, language level, British or American English and so on.

(b) the effectiveness of various methods of presenting information in monolingual dictionaries, for example for decoding (translating from L2 into L1 and comprehension of L2) and encoding (free expression, but not L1-L2 translation without using a bilingual dictionary first).

(c) the effectiveness of various methods of presenting information in bilingual dictionaries, for example for decoding (translating from L2 into L1 and comprehension of L2) and encoding (translating L1-L2, and free expression in L2).

In the samples of actual tests, the following items were included as test items:

(1) to test whether the user can handle polysemy.
(2) to test comprehension of parts of speech.
(3) to test comprehension of dictionary conventions and grammatical terms.
(4) to compare the effectiveness of various ways of giving the same information.
(5) to test the effectiveness of various types of dictionaries and ways of presenting information for purposes of vocabulary selection.
(6) to discover where the user expects to find such things as compound nouns, nominal and verbal idioms, prepositional and phrasal verbs, plural senses of words, etc.

Although this was not a complete list of the items to
be tested, it gives us a good idea about what they thought should be in the dictionary using skills test. Some of them (2, 3, and 6) are to be done without a dictionary, and others with a dictionary.

Okayama (1985) proposed to define the dictionary using skills as the combination of the following components:

(1) REFERENCE: a) knowledge of an alphabetical order   
   b) reference speed

(2) ANALOGY: a) descrimination of parts of speech   
   b) analogy of meanings

(3) COMPREHENSION OF DICTIONARY CONVENTIONS:   
   a) phonetic signs: i) stress system   
      ii) sound system   
   b) parts of speech labels   
   c) word forms   
   d) countable vs. uncountable   
   e) derivative forms   
   f) synonyms and antonyms

(4) GETTING INFORMATION IN A DICTIONARY   
   a) usage   
   b) social/cultural backgrounds

Checking understanding of an alphabetical order is necessary at least for the Japanese learners of English, for, as Whitcut (1986:112) said, "If the learner's L1 uses a different alphabet, even alphabetisation is quite a hurdle to be overcome." Okayama group also claimed that reference speed is important in assessing the reference skills. This is certainly true because we cannot call a person a skilled user when he spends ten minutes looking up one word in a dictionary. However, we have to be careful about the possibility of testing only one
aspect of the reference skills. There may be those users who are not used to using dictionaries against time. For such users the results may not reflect their true abilities.

Let us summarize the test items in the two sample tests above:

<table>
<thead>
<tr>
<th></th>
<th>Atkins et al.</th>
<th>Okayama</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alphabetisation</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Finding the word on the page</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Reference speed</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Sound system</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Stress system</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Parts of speech</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Grammatical terms</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Polysemy</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Vocabulary selection</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Word forms, inflections</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Countable vs. uncountable nouns</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Derivative forms</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Synonyms and antonyms</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>L2-L1 translation</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>L1-L2 translation</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Compounds, phrasal verbs, etc.</td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>

Okayama (1985) did not include the translation task with a dictionary because they were planning to make another set of test which would assess the use of dictionaries for L1/L2 or L2/L1 translation. While Okayama listed six items to be tested for dictionary conventions, Atkins did not specify which of the conventions should be tested. With a limited number of samples in this pilot study, we cannot understand what their assessment test is like exactly. They said that they were going to make a further interim report, which is to clarify their idea better.
What seems to be lacking in their selection of the test items is the examination of the natural process of dictionary look-up. For example, Scholfield (1982) describes the process of dictionary look-up as follows:

1. Locate the word(s) or phrase you don't understand.
2. If the unknown is inflected, remove the inflections to recover the form to look up.
3. Search for the unknown in the alphabetic list.
4. If you can't find at least one main entry for the unknown, try the following procedures. (Then he showed us the case for compounds, stems with suffixes, irregularly inflected forms, and an addendum.)
5. Understand the definition and integrate it into the context where the unknown was met.
6. If none of the senses entered seems to fit, attempt to infer one that does from the senses you have. If more than one fits, seek further contextual clues in the source text to disambiguate.

This may not be a perfect picture of the look-up process, but it still tells us how complicated the steps of looking a word up in a dictionary is, and how much active participation is required for the user. Do the assessment tests above, then, truly test this whole process of dictionary look-up? The answer is negative. The first step requires the ability to locate the word you do not understand, which is easier said than done. The skilled reader does not look up every word he/she does not know. Rather he/she will choose the unknown word which is most relevant to the content of the text and look it up.
first. This kind of intricate skills cannot be assessed in the above types of tests. The steps (5) and (6) are also very important skills. Especially the sixth step involves understanding of the overall meanings of the unknown, which helps infer the suitable meaning in the given source text. However, the user has difficulty looking through all the definitions of the word and usually stops at the first most likely definition (Tono 1984). The two tests above cannot assess this skill of inferring the meaning from the given list of definitions.

Another problem of the test made by the Okayama group is that they did not distinguish the two types of dictionary use; for recognition and for production. For instance, most of the test items such as testing comprehension of dictionary conventions (where to find inflected forms; where to find derivative forms; where to find antonyms and synonyms; etc.) are not very natural in assessing reference skills for comprehension. They are instead useful for production purposes. The mixture of the two different uses of dictionaries makes the test rather unclear about its objectives.

So far, we have looked at the contents of the two assessment tests and discussed the components of the dictionary using skills test. In the next chapter, the Okayama test being administered, the results of the tests will be reported and the implication discussed.

4. THE RESULTS OF THE TEST MADE BY THE OKAYAMA GROUP
4.1. SUBJECTS

Okayama (1985) reported that they administered the
dictionary using skills tests to 1,055 1st-year senior high school students in 12 different high schools in Okayama in September, 1984. The raw scores for each individuals could not be obtained for this paper. Therefore, in order to do more statistical analyses, this same set of tests was administered to 76 third grade students in Tokyo Metropolitan College of Aeronautical Engineering.

The Okayama's subjects included 536 students of high schools, 159 students of technical high schools, 271 students of commercial high schools, and 89 students of a home economics department. Some of the high schools were very high in the students' academic levels and many went to college after graduation, while the levels of most of the commercial and technical high schools were rather low. On the other hand, the students of Tokyo Metropolitan College of Aeronautical Engineering were mediocre in their English proficiency. Although it is called "college", it is actually a five-year school (equivalent to the 10th grade through the sophomore in college), where students learn to be mechanics or specialists in technology.

4.2. PROCEDURE

In the case of Okayama (1985), the students were asked to take the tests by listening to the tape which had all the instructions necessary for the tests. The questions [1] to [3] had to be answered without a dictionary, and the rest ([4] to [7]) with a dictionary. The time taken for the tests was 50 minutes.

The subjects at Tokyo Metropolitan College of
Aeronautical Engineering were all asked to take the tests during their class hours as a partial requirement of the course. We did not use a tape recorder, but the time limit for each test item was strictly kept and the tests were given within 50 minutes.

4.3. RESULTS AND DISCUSSION

Table 1 shows the results of the Dictionary Using Skills Tests made by the Okayama group. It shows the percentage of correct answers in each test item. The numbers before each test category refers to the question numbers. (For the test sample, please look at the appendix.) The two different results are shown, one reported in Okayama (1985) and the other for this present study. The two results are very similar in many senses. In spite of the age differences between the two groups, they performed in a similar way in many cases (e.g. the sound and stress system, parts of speech and their labels in dictionaries, inflections, synonyms and antonyms). The mean of the percentages of correct answers of the two groups are exactly the same (in both cases, the mean is 62).

In the question [41], the subjects were asked to look up as many as 12 words to choose the right definitions within three minutes. None of them looked up all the words within the time limit. Most of the students tried hard with the first set of words (1-6) but got discouraged with the second set (7-12). As was pointed out earlier, in many cases, the subjects were asked to use dictionaries in a rather unnatural way. There are not so many cases in real life situations, where you have to look up as many words as you
<table>
<thead>
<tr>
<th>TEST ITEMS</th>
<th>OKAYAMA(1985)</th>
<th>TONO</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1] SOUND SYSTEM</td>
<td>72%</td>
<td>72%</td>
</tr>
<tr>
<td>[2] ALPHABETISATION</td>
<td>89%</td>
<td>98%</td>
</tr>
<tr>
<td>[3a] PARTS OF SPEECH</td>
<td>54%</td>
<td>54%</td>
</tr>
<tr>
<td>[3b] ANALOGY OF MEANINGS A</td>
<td>56%</td>
<td>71%</td>
</tr>
<tr>
<td>[3c]</td>
<td>68%</td>
<td>67%</td>
</tr>
<tr>
<td>1</td>
<td>86%</td>
<td>89%</td>
</tr>
<tr>
<td>2</td>
<td>86%</td>
<td>81%</td>
</tr>
<tr>
<td>3</td>
<td>60%</td>
<td>76%</td>
</tr>
<tr>
<td>4</td>
<td>62%</td>
<td>71%</td>
</tr>
<tr>
<td>5</td>
<td>45%</td>
<td>-63%</td>
</tr>
<tr>
<td>6</td>
<td>31%</td>
<td>63%</td>
</tr>
<tr>
<td>7</td>
<td>12%</td>
<td>13%</td>
</tr>
<tr>
<td>8</td>
<td>9%</td>
<td>8%</td>
</tr>
<tr>
<td>9</td>
<td>8%</td>
<td>8%</td>
</tr>
<tr>
<td>10</td>
<td>7%</td>
<td>5%</td>
</tr>
<tr>
<td>11</td>
<td>4%</td>
<td>0%</td>
</tr>
<tr>
<td>12</td>
<td>8%</td>
<td>10%</td>
</tr>
<tr>
<td>[5] STRESS SYSTEM</td>
<td>51%</td>
<td>50%</td>
</tr>
<tr>
<td>[6a] PARTS OF SPEECH LABELS</td>
<td>67%</td>
<td>71%</td>
</tr>
<tr>
<td>[6b] INFLECTIONS</td>
<td>77%</td>
<td>78%</td>
</tr>
<tr>
<td>[6c] UN/COUNTABLE BOUNDS</td>
<td>67%</td>
<td>53%</td>
</tr>
<tr>
<td>[6d] DERIVATIVES</td>
<td>55%</td>
<td>46%</td>
</tr>
<tr>
<td>[6e] SYNONYMS/ANTONYMS</td>
<td>59%</td>
<td>61%</td>
</tr>
<tr>
<td>[7a] USAGE</td>
<td>46%</td>
<td>64%</td>
</tr>
<tr>
<td>[7b] SOCIAL/CULTURAL BACKGROUND</td>
<td>68%</td>
<td>90%</td>
</tr>
</tbody>
</table>
Table 2  The descriptive statistics of the Dictionary Using Skills Tests

<table>
<thead>
<tr>
<th>N</th>
<th>MEAN (RAW SCORES)</th>
<th>MEAN (X)</th>
<th>S. D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>76</td>
<td>43.8</td>
<td>62.5</td>
<td>7.5</td>
</tr>
</tbody>
</table>

Table 3  The correlations of the test items in the Dictionary Using Skills Tests

<table>
<thead>
<tr>
<th></th>
<th>[1]</th>
<th>[2]</th>
<th>[3a]</th>
<th>[3b]</th>
<th>[3c]</th>
<th>[4]</th>
<th>[5]</th>
<th>[6a]</th>
<th>[6b]</th>
<th>[6c]</th>
<th>[6d]</th>
<th>[6e]</th>
<th>[7a]</th>
<th>[7b]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1]</td>
<td>.18</td>
<td>.00</td>
<td>.06</td>
<td>.06</td>
<td>.18</td>
<td>.12</td>
<td>.13</td>
<td>-.13</td>
<td>-.12</td>
<td>.13</td>
<td>.25</td>
<td>.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[2]</td>
<td>.18</td>
<td>.12</td>
<td>.07</td>
<td>-.54</td>
<td>-.02</td>
<td>.00</td>
<td>-.03</td>
<td>.41</td>
<td>.04</td>
<td>.04</td>
<td>.11</td>
<td>.33</td>
<td>.15</td>
<td></td>
</tr>
<tr>
<td>[3a]</td>
<td>.00</td>
<td>.12</td>
<td>-.04</td>
<td>.11</td>
<td>.30</td>
<td>.22</td>
<td>.13</td>
<td>.21</td>
<td>.45</td>
<td>-.09</td>
<td>.12</td>
<td>.26</td>
<td>.18</td>
<td></td>
</tr>
<tr>
<td>[3b]</td>
<td>.06</td>
<td>.07</td>
<td>-.04</td>
<td>.04</td>
<td>-.19</td>
<td>-.14</td>
<td>-.11</td>
<td>-.36</td>
<td>-.09</td>
<td>-.03</td>
<td>-.05</td>
<td>.18</td>
<td>-.18</td>
<td></td>
</tr>
<tr>
<td>[3c]</td>
<td>.06</td>
<td>.54</td>
<td>.11</td>
<td>.04</td>
<td>.08</td>
<td>.03</td>
<td>-.06</td>
<td>.13</td>
<td>-.03</td>
<td>-.02</td>
<td>.14</td>
<td>.33</td>
<td>.21</td>
<td></td>
</tr>
<tr>
<td>[4]</td>
<td>.01</td>
<td>-0.2</td>
<td>.30</td>
<td>-.19</td>
<td>.08</td>
<td>-.45</td>
<td>.44</td>
<td>.27</td>
<td>.33</td>
<td>-.25</td>
<td>.31</td>
<td>.34</td>
<td>.25</td>
<td></td>
</tr>
<tr>
<td>[5]</td>
<td>.18</td>
<td>.00</td>
<td>.22</td>
<td>-.14</td>
<td>.03</td>
<td>.45</td>
<td>.40</td>
<td>.45</td>
<td>.44</td>
<td>.38</td>
<td>.28</td>
<td>.30</td>
<td>.16</td>
<td></td>
</tr>
<tr>
<td>[6a]</td>
<td>.12</td>
<td>-.03</td>
<td>.13</td>
<td>-.11</td>
<td>-.06</td>
<td>.44</td>
<td>.40</td>
<td>.37</td>
<td>.23</td>
<td>.22</td>
<td>.28</td>
<td>.02</td>
<td>.17</td>
<td></td>
</tr>
<tr>
<td>[6b]</td>
<td>.13</td>
<td>.41</td>
<td>.21</td>
<td>-.36</td>
<td>.13</td>
<td>.27</td>
<td>.45</td>
<td>.37</td>
<td>.31</td>
<td>.31</td>
<td>.30</td>
<td>.21</td>
<td>.39</td>
<td></td>
</tr>
<tr>
<td>[6c]</td>
<td>-.13</td>
<td>.04</td>
<td>.45</td>
<td>-.09</td>
<td>-.03</td>
<td>.33</td>
<td>.44</td>
<td>.23</td>
<td>.31</td>
<td>.34</td>
<td>.17</td>
<td>.11</td>
<td>.12</td>
<td></td>
</tr>
<tr>
<td>[6d]</td>
<td>-.12</td>
<td>.04</td>
<td>-.09</td>
<td>-.03</td>
<td>-.02</td>
<td>.25</td>
<td>.38</td>
<td>.22</td>
<td>.31</td>
<td>.34</td>
<td>.37</td>
<td>.03</td>
<td>.18</td>
<td></td>
</tr>
<tr>
<td>[6e]</td>
<td>.13</td>
<td>.11</td>
<td>.12</td>
<td>-.05</td>
<td>.14</td>
<td>.31</td>
<td>.28</td>
<td>.28</td>
<td>.30</td>
<td>.17</td>
<td>.37</td>
<td>.20</td>
<td>.54</td>
<td></td>
</tr>
<tr>
<td>[7a]</td>
<td>.25</td>
<td>.33</td>
<td>.26</td>
<td>.18</td>
<td>.33</td>
<td>.34</td>
<td>.30</td>
<td>.02</td>
<td>.21</td>
<td>.11</td>
<td>.03</td>
<td>.20</td>
<td>.17</td>
<td></td>
</tr>
<tr>
<td>[7b]</td>
<td>.28</td>
<td>.15</td>
<td>.18</td>
<td>-.18</td>
<td>.21</td>
<td>.25</td>
<td>.16</td>
<td>.39</td>
<td>.12</td>
<td>.18</td>
<td>.54</td>
<td>.17</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(*) = significant at .01 level)
can in three minutes. Or they do not usually refer to
dictionaries to know what parts of speech a certain
word belongs to. They can tell from the context and
the sentence structure. If you really want to know
whether the user understands the symbols for the
parts of speech, it would be better to ask what they
are directly, instead of asking them to look words up
just for that purpose, which is rather unnatural.

Since the data presented in Okayama (1985) lacks
some other descriptive and inferential statistics,
the data taken for this research was statistically
analyzed to supply more data. Table 2 shows the basic
descriptive statistics and Table 3 presents the
results of Pearson $r$ between each pair of the test
items.

Table 3 indicates that there is a rather low cor-
relation between each test item. We can think of some
reasons for this low correlation. First, if we take
this data positively, we can claim that each of the
reference skills is quite independent from each
other. In other words, for instance, being skilled in
looking up many words in a limited time is quite dif-
ferent from being skilled in finding whether or not a
certain noun is countable in a dictionary. From this
viewpoint, the data indicates that the skills tested
here are all independent abilities and therefore do
not affect one another. If this is really the case,
each skill can be acquired or taught independently.

However, it is doubtful whether this is true with
all the reference skills. For example, there are some
cases in which no positive correlation was found
where there should be one. The Okayama group said
that the ability to guess the parts of speech and the
meanings of the words from the context of the source text was essential for using dictionaries effectively (p. 113). This is a well-grounded statement and thus they intended to test this very skill in the question [3]. The results, however, showed that no positive relation was found between this ability and other dictionary using skills except for one ([3a] and [6c]). Here we cannot just conclude that there is no significant relation there, but rather have to doubt the validity of the test items. There may be some skills which are quite independent from each other and thus acquired that way, but for the rest of the skills, there should be some relations among the skills and this Okayama version could not give us a clear picture of how those skills were interrelated. Further clarification is needed in this area.

Time limitation was also another factor which made the look-up process rather unnatural. It is true that we need a certain amount of speed in looking up, but the users are not often aware of the time taken for look-up. There is a danger in this kind of test that the subjects were forced to look up words in a limited time and that the process of look-up became different from what it really was.

Finally, the types of dictionaries they used in the tests were to be considered in order to evaluate their skills. For instance, some dictionaries present parts of speech in English, e.g. n., adj., conj., and the like, while the learner's dictionaries published recently all use Japanese symbols such as [名], [形], and [接]. The results will be very different if they do not know the English abbreviations. New learner's dictionaries contain derivatives under the
main entries of the stem words but old ones do not usually. So if you have old dictionaries, you have to look for derivatives around the entries of root words. This is another example which makes difference in their look-up skills. The Okayama test was made in consideration of several popular bilingual learner's dictionaries, most of which had this advanced feature. Further investigation is needed in this relation of dictionary layout with the user’s using skills.

5. CONCLUSION

So far we have discussed how we could assess the dictionary using skills in a most valid way and examined the two assessment tests, one by Atkins et al. and the other by the Okayama group. Although the one by Atkins et al. is still a pilot study and we have not been able to examine its details, these two tests have lent us good insight into the optimal dictionary using skills test.

The most significant problem of making a good test is to clearly see what you are going to test. Here we still wonder what exactly the dictionary using skills are. The Okayama group interpreted that they were the combination of various sub-skills and thus made a discrete-point test. The results, however, made us think of the importance of the opposite approach.

Further research on how to increase the validity of the test items is definitely important. Seeking for the improvement of the Okayama test by reconsidering the test items and narrowing down the target or purpose of the test is one way and, at the same time, a different approach to assess the reference
skills must be explored. That is, as we have two kinds of tests, objective tests and essay-type tests in measuring things, we should also try to make an essay-type test for the dictionary using skills. With this approach, we can deal with the reference skills as a whole, as an entire process. The comparison of the two different types of the tests will throw light on different aspects of the dictionary using skills in future.

REFERENCES


Appendix The Okayama’s Dictionary Using Skills Test (Sample)

英語辞書力診断テスト(1)

岡山県高等学校教育研究会英語部会編

次の問題は、辞書がどれだけ引けるか試す問題です。テープの指示に従って、1～4は辞書を引かずに答え、裏面の5～8は辞書を引いて答えない。

解答用紙記入上の注意）をよく読み、答えはすべて解答用紙乙欄の1～70に従ってマークしなさい。

1 発音された語を発音記号で示すとすれば、それぞれ0)または1)のどちらですか。（2分）
1) (0) [kap] (1) [kæp]
2) (0) [hæŋ] (1) [hæŋ]
3) (0) [fɪp] (1) [fɪp]
4) (0) [kaut] (1) [kɑːt]
5) (0) [swɪtʃ] (1) [swɪtʃ]
6) (0) [stː] (1) [stː]
7) (0) [kwɪstʃən] (1) [kwɪstʃən]
8) (0) [ˈfɔːrər] (1) [ˈfɔːrər]

2 次の(A), (B)の問いに答えなさい。（2分）
(A) 次の5, 10について、アルファベット順に並んでいる組を、それぞれ0)～3)から選びなさい。
9) (0) r–q–p (1) q–p–r (2) p–q–r (3) p–r–q
10) (0) w–v–u (1) v–u–w (2) u–v–w (3) u–w–v

(B) 次の11, 12について、辞書に出てくると思われる順に並んでいる組を、それぞれ0)～3)から選びなさい。
11) (0) noise —— mistake —— junior —— today
(1) junior —— mistake —— noise —— today
(2) junior —— mistake —— today —— noise
(3) mistake —— noise —— junior —— today
12) (0) ship —— shirt —— skirt —— sky
(1) sky —— shirt —— ship —— skirt
(2) sky —— shirt —— skirt —— ship
(3) ship —— skirt —— sky —— shirt
次の(A)～(C)の問いに答えなさい。

(A) 次の下線部の品詞を推察し、それぞれ(0)〜(5)から選びなさい。(3分)

13. Please give me some water.
14. Her job was to water the horses.
15. Hanako studies English very hard.
16. It was a hard question to answer.
17. Let's go swimming after school.
18. She arrived after he left.
19. What a kind girl Hiroko is!
20. What kind of flowers are these?

(0) 名詞 (1) 形容詞 (2) 副詞 (3) 動詞 (4) 前置詞 (5) 接続詞

(B) 次の下線部の動詞の意味を推察し、それぞれ(0)〜(3)から選びなさい。(2分)

21. A fine snow was falling outside.
22. I can't stand the cold.
23. He has a right to say it.

(0) 立てる (1) 我慢する (2) おどる (3) 合格する

(C) 次の下線部の動詞が同じ意味で用いられている文を、それぞれ(0)〜(2)から選びなさい。(3分)

24. Does this meet your hopes?
25. I have a pen-friend who lives in Australia.
26. She carries her baby on her back.

(0) 明るい (1) 元気な (2) 細かい (3) 立派な
(0) 言う (1) 勝つ (2) 延ばす (3) 正しい
(0) 働く (1) 正義 (2) 権利 (3) 正しい

24. I am going to meet him at the station.
25. He lives alone in the country.
26. She carries her baby on her back.

(1) Two lives were lost in the accident.
25. I am reading a book about the lives of great men.
26. The store stands back from the street.

(2) I am glad to meet you.
25. He lives alone in the country.
26. He turned his back to her.
（次のⅣ～Ⅵは辞書を引いて答えなさい。）

Ⅳ 次の(A)、(B)欄の語の意味を、それぞれ(0)～(6)から選びなさい。（3分）
(A) 27 quinsy 28 hydrangea 29 autism 30 sediment
31 linguistics 32 chrysalis
(0) アジサイ (1) 言語学 (2) きなぎ (3) 自閉症 (4) 沈殿物 (5) ヘンナ
(6) へんとうせん炎

(B) 33 traitor 34 malady 35 falconry 36 wrath
37 optimism 38 veto
(0) 拒否権 (1) 暴怒 (2) 食欲 (3) タカ特 (4) 反逆児 (5) 違背
(6) 楽天主義

Ⅴ 次の語の最も強いアクセントがある位置を、それぞれ(0)～(4)から選びなさい。（解答用紙乙
面の39からマークすることを確認しなさい。）（2分）
39 col-lapse 40 guair-an-tee 41 fra-ter-ni-ty 42 an-thro-pol-o-gy
(0) (1) (0) (1) (2) (0) (1) (2)(3)(4)

Ⅵ 次の(A)～(E)の問いに答えなさい。
(A) 次の語の品詞をそれぞれ(0)～(5)から選びなさい。（2分）
43 sacred 44 inquiry 45 whereas 46 maybe
(0) 名詞 (1) 形容詞 (2) 副詞 (3) 動詞 (4) 接続詞 (5) 前置詞

(B) 次の語の指示された形をそれぞれ(0)～(2)から選びなさい。（3分）
47 seek （過去形） (0) seek (1) seeked (2) sought
48 admit （～ing 形） (0) admitting (1) admitting (2) admeting
49 ripe （最高級） (0) most ripe (1) ripest (2) most ripest
50 crisis （複数形） (0) crises (1) criseses (2) criseses

(C) 次の語は、(0)数えられる名詞、または(1)数えられない名詞のどちらですか。（2分）
51 kitten 52 privacy 53 soap 54 ladder

(D) 次の語の指示された形を、それぞれ(0)～(2)から選びなさい。（3分）
55 strong （名詞形） (0) strongness (1) storongment (2) strength
56 fortune （形容詞形） (0) fortunate (1) fortunable (2) fortunative
57 expansion （動詞形） (0) expanse (1) expant (2) expand
58 scarce （動詞形） (0) scarcity (1) scatty (2) scarcely
(E) 次の語の指示された語を、それぞれ①〜②から選びなさい。（3分）
- danger （同意語） ① risk ② safety ③ anger
- bare （同意語） ① severe ② naked ③ bear
- artificial （反意語） ① practical ② official ③ natural
- minority （反意語） ① height ② maturity ③ majority

Ⅶ 次の(A), (B)の問いに答えなさい。
(A) 次の英文を完成するためには、どの語が最も適切ですか。それぞれ①〜②から選びなさい。 （4分）
- To eat mochi on New Year's Day is a Japanese ( )
  (正月におもちを食べるのが日本の習慣です。)
  ① habit ② custom ③ routine
- The bus ( ) to our school is 100 yen.
  (私たちの学校までのバス料金は100円です。)
  ① fare ② cost ③ price
- We ( ) this apartment at 40,000 yen a month.
  (私たちはこのアパートを1か月4万円で借りている。)
  ① borrow ② rent ③ use
- He opened the window and put his ( ) out of it.
  (彼は窓を開けて、頭を出した。)
  ① face ② neck ③ head

(B) 次の問いに対する答えを、それぞれ①〜②から選びなさい。（3分）
- generation は約何年にあたりますか。
  ① 約20〜30年 ② 約50〜60年 ③ 約75〜100年
- brunch はもともとの組み合わせからできた語ですか。
  ① brother + punch ② breakfast + lunch ③ brush + branch
- the first floor はイギリス英語では何階ですか。
  ① 地下一階 ② 一階 ③ 二階
- foot は下の絵のどの部分ですか。
  ① ひざからくる部分まで
  ② くらぶしか下