

## **Holistic Evaluation of English Ability By Measuring Each Skill: A Pilot Study**

Keiko SATO

### **INTRODUCTION**

There are four types of language tests depending on their purposes; achievement test, which assesses how much knowledge is learned during a definite period; proficiency test, which assesses the ability of a learner at a definite period in time; diagnostic test, whose purpose is to get as much information as possible about a learner's errors to analyze for the remedial teaching; and aptitude test, which measures ability and predicts if a learner can learn a given subject, such as mathematics or a foreign language (Kakita and Aoki, 1985).

Furthermore, each of these has subcategories, with some variation according to test specialists. The so-called standard test is developed by an institution for a large number of learners with different educational backgrounds and a wide-range of ability. There are tests made by teachers of a certain subject for their students so that they can get in touch with them in class and become aware of their learning processes (David Harris, translated by Ohtomo, 1972). Few people would disagree that better instruction can be given if an instructor knows his or her student's level of attainment and proficiency in detail.

### **Purpose**

Our primary purpose was to develop a test accurately to measure English language ability and establish correlations between each of four skills — hearing, speaking, reading, and writing — and their integrative language ability. A large number of experiments will be required, however, as pilot tests, because several kinds of test must be examined for their reliability, validity, scorability, economy, and ease of administration.

This study is a pilot study in which the discrete point approach in tests is expected to show the continuous result in and strong relationship with the measurement of global language ability. The hypotheses tested in this experiment are listed below:

### **Hypotheses**

1. Language ability is developed as a whole, so late learners are deficient in every skill.
2. One or more of the four skills will be less developed when compared with the other skills.

## PROCEDURE

### Test Design

Language is a complex multifaceted skill, involving the process of encoding (speaking and writing) and decoding (hearing and reading). These so-called 'four skill' all require two important elements; vocabulary and grammar. For this reason, unique subtests should be designed for investigating the six dimensions mentioned above for each learner's English language ability at any given time. If someone has an all-around skill in every aspect of English activity, his or her score can be represented as figure 1. If someone has a deficiency in some skill, it will cause the figure to be an indented hexangle.

The measured language elements and skills:

1. Vocabulary
2. Grammar
3. Hearing
4. Comprehension
5. Writing
6. Pronunciation

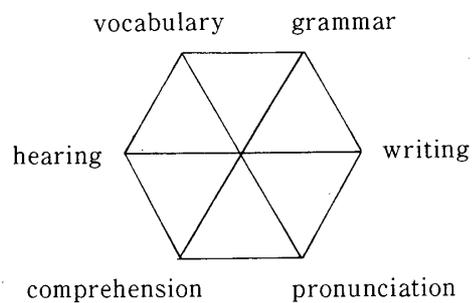


Figure 1

The vocabulary test used was developed from the items used at the vocabulary competition held in 1986 by the publisher Kenkyusha. This test covered almost all aspects of vocabulary. Six categories were used for our scales; spelling, pronunciation, stress, meaning, synonym/antonym and idiomatic expression.

As the grammar test, ten categories of grammatical rules with three questions each were selected. These categories are concordance (or agreement), tense, subjunctive mood, auxiliary, voice, transitive and intransitive, infinitive, participle, gerund and others. The form of questioning used is cloze by alternative choice. This form is easier than multiple choice or completion. The reason for adopting this is that it is intended to examine only basic knowledge of English grammar.

As the hearing test, the entire 40 questions of the JACET Basic Listening Comprehension Test was given to the subjects. The complete version of this test is prohibited from public presentation by JACET. This study was conducted in the Nagoya Women's University LL room. The test took thirty-five minutes to conduct following the manual direction, and scoring was left to JACET Test specialists.<sup>1</sup>

As the comprehension test, cloze test was constructed that focussed separately on content words and function words. Material used in the test was taken from "The Birth of Jesus" in a college textbook published by Asahi Shuppan, "The Examination Hell in Britain," and "The Post Office" in *Close Test and English Educa-*

*tion*. Those three materials were selected respectively as the categories of narrative, essay, and conversation in style.<sup>2</sup>

As the writing test, a paragraph writing entitled "A Famous Place to Visit in Japan" was written during the subject's term-end examination of composition. The following aspects were inspected as the measured elements; length of the whole paragraph by words, number of sentences, number of T-units, number of complex sentences, error scarcity and logical sequences of content. The last one was evaluated impressionistically by three levels.<sup>3</sup>

As the substituted test for speaking, students read a passage aloud and answered five oral questions. This was recorded onto tape and evaluated for 6 aspects of speech quality; vowels and consonants, sequences of sounds, intonation, rhythm, sentence stress, and fluency. For the conveniences of evaluation, the places in the passage to be analyzed were decided beforehand. The test passage was selected from the past Eiken 3rd grade tests containing suitable sounds for our investigation.<sup>4</sup>

## Subjects

The tests designed as described above were given to 50 students—all from one class—majoring in English. All the subjects were female, aged 18 to 19. The first test was given just after they entered college. The last test was given 6 months after the first. All subjects are supposed to be almost at the same level of English language ability.

## Scoring Method

The tests used in this study cannot be compared because of differences between them. That is why neither the raw scores of the subject on each category nor their percentage can be compared. Therefore, the standard scores (SS) for each subject and for each English category were calculated from the raw score, the mean, and the standard deviation (SD). The standard score shows a subject's deviation from the average score of the corpus, the relative position in the same corpus. The formula for SS is as follows:<sup>5</sup>

$$SS = ((X - \bar{X}) / SD) \times 10 + 50$$

Here  $X$  represents a subject's raw score and  $\bar{X}$  the average raw score (mean) of the corpus. The mean of SS's in the corpus must be 50 and the new SD from SS must be 10.

Standard scores were divided in 5 categories by the grading method used by JACET. The subjects in the highest and lowest categories of English ability were selected in order to compare ability in each category with that in other categories. To compare different subject's ability, graphs were constructed showing this com-

parison.

rank	SS	distrubution	%
S	over 65	Highest	7
A	55 to 64	Second highest	24
B	45 to 54	Middle	38
C	35 to 44	Second lowest	24
D	bellow 35	Lowest	7

Chart 1

## RESULT and DIDCUSSION

### Vocabulary

The two highest scoring subjects consistently obtained the highest score in every category of English ability, except hearing. The other two high scoring subjects did not show the same consistency. This shows that their ability over the different levels of ability are not the same.

The four lowest scoring subjects scored consistently in the lowest level of every category of ability. The one exception to this is that two of them scored in the top level of writing category. This is probably true because the questions of the hearing test contained no difficult words and during the writing test subjects could avoid the use of words they did not know.

Relation of Vocabulary and Skills in English  
-From Vocabulary Amount

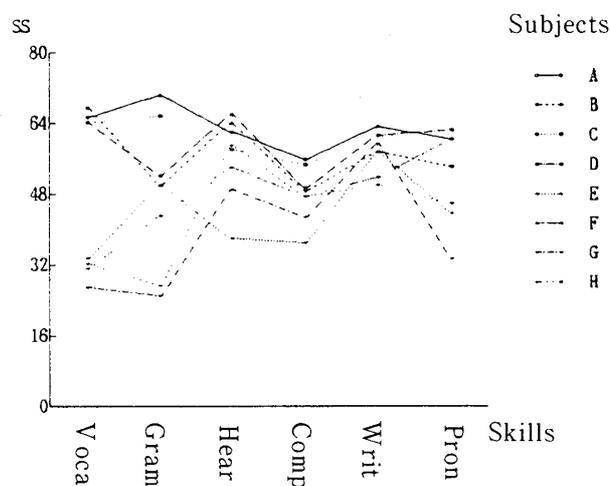


Figure 2

### Grammar

As in the case above, the three highest and four lowest scoring subjects in the category of grammar were isolated for making the graph indicating the interrelation of grammar knowledge and other four skills (Figure 3). Notice that the highest

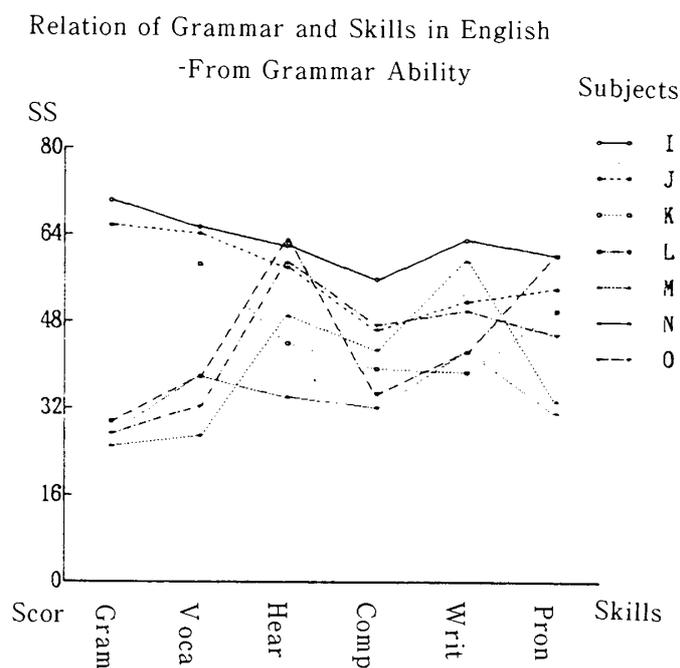


Figure 3

scoring subjects consistently obtained the top score for every category and that the lowest scoring subjects obtained inconsistent scores in the different categories. This finding shows mixed support for the two hypotheses. Hypothesis 1 is supported for advanced learners, while hypothesis 2 is supported for the less advanced learners. Particularly strong support was found for the relationship that grammar is associated with reading comprehension and writing. The relationship that grammar held with hearing and pronunciation was not so strong.

### Hearing

An examination of figure 4 shows that hearing score is not always related to scores in other areas of ability. Only one of the high scoring and two of the low scoring subjects maintained almost horizontal level in the other categories. In particular, two of the high scoring subjects are outstandingly good in hearing, but notably bad in writing. Hearing ability, on the other hand, seems consistent with pronunciation ability. This point will be restated to later.

### Reading Comprehension

Figure 5 shows a clear relationship between reading comprehension and other language abilities. The one exceptional subject to this is a surprising counter example. The subject has acquired a considerably large vocabulary and good grammatical knowledge, while her other English skills have not yet reached the average level.

Reading comprehension requires the use of many skills. For example, in order to understand a passage, readers must think of ideas depending on their current linguistic ability. This includes their vocabulary acquired, grammatical and structural knowledge, and the context of the passage. In order to place a proper word into a blank in the passage or choose a more suitable word among given items, the reader will have to make logical inferences and use contextual clues determined by their accuracy and their speed of perception of written symbols. This seems to be why the sample indicating a high ability in reading comprehension surpasses the other subjects in other categories, and vice versa.

## Writing

As with reading comprehension, writing freely and spontaneously requires other skills such as language knowledge, skills, and their combinations.

In Figure 6, it is shown that good writing is associated with a large vocabulary and good knowledge of grammatical rules, though hearing and pronunciation are not related so strongly with writing.

Writing is a very complex task that involves many skills. The test of writing composition should be a test of the production of ideas and their organization in the advanced level. In the less advanced level, on the other hand, spelling and punctuation, paragraphing, appropriateness of vocabulary, accuracy of grammatical structures, proper length of sentences, frequency of complex clause use, and logical and informational sequences should mainly be assessed (Finocchiaro and Sako, 1983). Writing is, in fact, the production of an individual's integrated language skills.

Relation of Hearing Ability and Other Skills  
From Hearing Ability

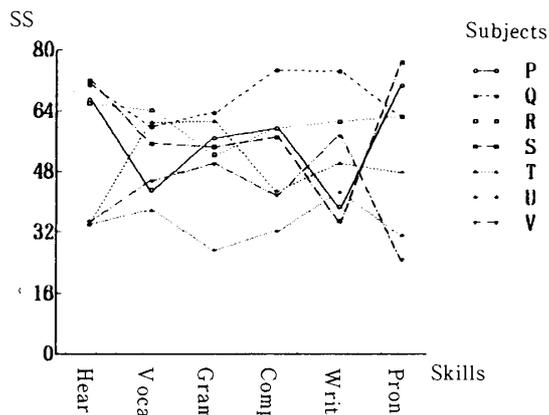


Figure 4

Relation of Reading Comprehension and Skills  
From Reading Passages

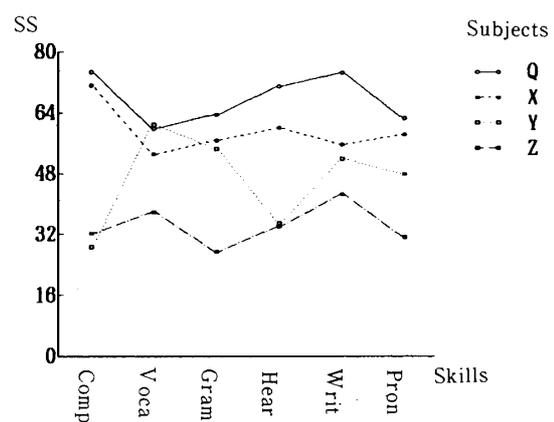


Figure 5

Relation of Spontaneous Writing and Skills  
From Writing

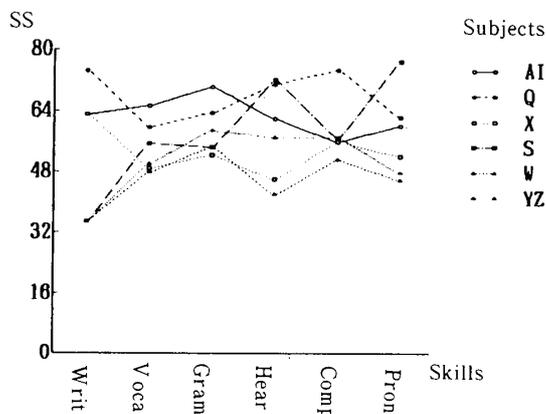


Figure 6

Relation of Pronunciation/Speaking and skills  
From Pronunciation

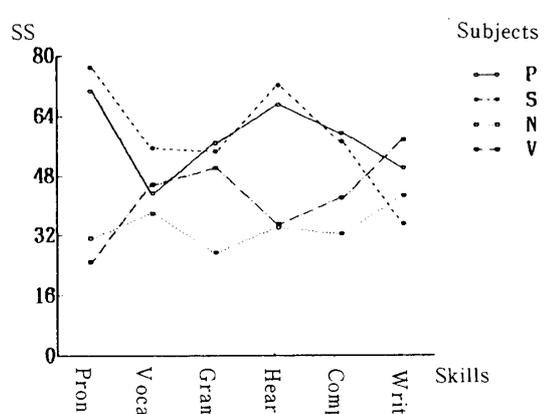


Figure 7

### Pronunciation/Speaking

Since real communication is consisted of speaker's and listener's roles, Eiken's interview test is a combination of oral product and aural comprehension. A subject's reading of a passage aloud without having seen it in advance reveals his or her present acquisition of the sound system.

Interchange between a subject and an interviewer in the question - and - answer type encourages an answer to understand the passage he or she has just read and to perceive the oral questions a questioner has asked before the former gives oral answers.

Accordingly, the result of an Eiken's interview test does not merely reflect how well the subjects can pronounce the words, can read the passage, or respond orally. It is a measure of the integrated ability of listening comprehension and oral production. However, unlike expressing his or her ideas freely in the upper class, in this Q - and - A type the questions are limited within the content of the reading passage and are therefore apt to be easy and simple ones. It might be better to call it a comprehensive speaking test rather than a speaking test.

Returning to our result, as shown in the next graph (Figure 7), the rank order of pronunciation/speaking and writing go almost in reverse. Pronunciation does not seem to be related to writing ability, rather it is actually concerned with an attitude toward sounds. People whose language superiority lies in hearing.

### CONCLUSION

The above results provide convincing support for the hypotheses. As seen from the next two graphes (Figures 8 and 9), which show the ability of representative subjects selected from each synthetic level, language ability has a strong tendency to develop with integrating several skills. Also less advanced learners have not reached the same level as advanced learners. The former are still on their way to the goal, and their present language is their interlanguage, which falls on a continuum. Learners who have reached the upper level are good in all the categories, while those who have not are poor in all the categories (Hypothesis 1). However, this case only can be said to a degree. There is some counter evidences seen especially in the samples of the low level. For this reason Hypothesis 2 cannot be rejected so far.

Relations of Skills in Each Level  
Representative Five Levels

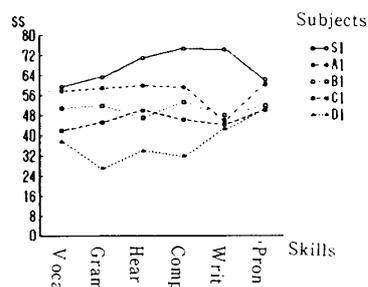


Figure 8

Relations of Skills in Each Level  
Representative Five Levels

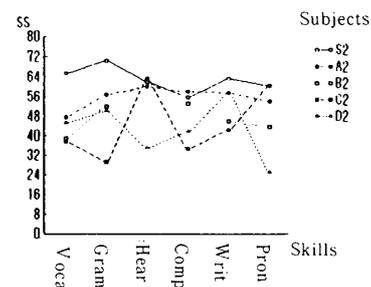


Figure 9

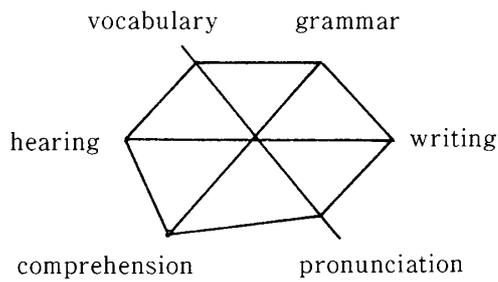


Figure 10 S1's Ability

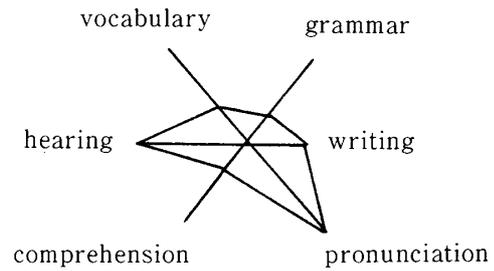


Figure 11 C2's Ability

### REMAINING PROBLEMS

In the course of this experiment, a number of problems were discovered with the hypotheses, though there were some supporting findings. First, there is a difference in difficulty among the tests across several categories of English ability. Secondly it is necessary to demonstrate the reliability of these tests before using them. Next, some of the tests used here were time-consuming, such as in the scoring method for writing, which made extensive use difficult. Finally, the use of high technology or a computer as the data-processing equipment, is important. With easy access to high technical equipment for processing data, it will be possible to make time for the preparation of tests and the contemplation of results and remediation.

### NOTES

1. They changed the raw scores by giving some different load value to each part before calculating out SS. We ignored this treatment of the scores of other categories.
2. Considering the result, the first two test passages might have been too difficult for the subjects. The content should be one that can be imagined from the words left behind in the passage.
3. It might have been a mistake to include the length of the whole paragraph. Some subjects took much time writing a paragraph, and others less time.  
See Gaies, S.J. 1977 for detail information of T-unit.
4. The following items were assessed; vowels [æ] in "glad" and "happy", [ə] in "were" and "mother"; consonants [θ] in "theater", [ð] in "mother", and [l] in "glad"; phrase stress in "būs driver" and "mōvie theater"; intonation in "understand him" and "understand her English"; breath position in "discovered that" and "told him that"; flow of sequences and fluency.
5. Manual for JACET Basic Listening Comprehension Test, Kaitakusha. 1990.

## REFERENCES

- Davies, Alan (ed.) 1968. Language Testing Symposium : A Psycholinguistic Approach, London: Oxford University Press.
- Duly, H. , Burt, M. and Krushen, S. 1982. Language Two, New York : Oxford University Press.
- Finocchiaro, M. and Sako, S. 1983. Foreign Language Testing : A Practical Approach, New York : Regents Publishing Company.
- Henning, Grant 1987. *A Guide to Langage Testing : Development, Evaluation, Research*, Cambridge, Massachusetts : Newbury House Publishers.
- JACET Study Group on the Evaluation of EFL/ESL Writing. "A study of the Measurement of EFL Writing : Can We Specify an Analytic Scoring Item Which High Correlation with Impressionistic Scoring." JACET Bulletin 20 (1989) : 17-36.
- Lado, Rovert 1961. Language Testing : The Construction and Use of Foreign Language Tests, London : Longmans.
- Madsen, Harold S. 1983. Techniques in Testing, London : Oxford University Press.
- Ohtomo, K. , Ashano H. , Hattori T. , and Yoshie M. , "Test-Free Person Measurement in Tests of English for Japanese Students." JACET Bulletin 20 (1989): 71-87.
- David P. Harris 1969.大友賢二訳註, 1981.『英語の測定と評価』東京, 英語教育協議会 (ELEC)
- 青木昭六編, 1985.『英語の評価論』東京, 大修館
- 吉田一衛編, 1984.『英語のリスニング』東京, 大修館
- 松村幹男編, 1984.『英語のリーディング』東京, 大修館
- 岡 秀夫編, 1984.『英語のスピーキング』東京, 大修館
- 沖原勝昭編, 1985.『英語のライティング』東京, 大修館
- 竹蓋幸生, 1989.『ヒアリングの指導システム』東京, 研究社
- 竹蓋幸生, 1982.『日本人英語の科学』東京, 研究社
- 竹蓋幸生, 1984.『ヒアリングの行動科学』東京, 研究社

APENDIX I  
VOCABULARY TEST

- ① 次の各組で、スペリングが正しいものを (a) ~ (b) の中からそれぞれ1つずつ選びなさい。
1. (a) calendar (b) calendor
  2. (a) umbrella (b) umbrellla
  3. (a) cigarette (b) cigarrette
  4. (a) penicilin (b) penicillin
  5. (a) prejudise (b) prejudice
- ② 次の各組で、太字で示された単語の意味として、最もふさわしい日本語を (a) ~ (d) の中からそれぞれ1つずつ選びなさい。
56. congratulato  
(a) 分類する (b) 共存する (c) つなぐ (d) 祝う
  57. turiff  
(a) 貿易 (b) 芝生 (c) 関税 (d) 税関
  58. indignant  
(a) 立派な (b) 無関心の (c) 憤慨した (d) 必要不可欠の
  59. ramp  
(a) (立体交差などの) ランプ, 斜道 (b) 灯火 (c) 尻, (牛の) 尻肉 (d) こぶ
  60. opportunist  
(a) 悲観主義者 (b) 日和見主義者 (c) 楽観主義者 (d) 物質主義者
- ③ 次の各組で、太字で示された日本語に最もふさわしい単語を (a) ~ (d) の中からそれぞれ1つずつ選びなさい。
126. 接着剤  
(a) detergent (b) adhesive (c) fertilizer (d) vaseline
  127. 宇宙遊泳  
(a) spacewalk (b) spaceswimming (c) spacecourse (d) space - exploration
  128. ブランコ  
(a) wave (b) stilts (c) swing (d) pinwheel
  129. 八角形  
(a) rectangle (b) hexagon (c) octagon (d) pentagon
  130. 文房具  
(a) stationery (b) utensils (c) furnishings (d) chemicals
- ④ 次の各組で、( ) に入れるのに最もふさわしいものを (a) ~ (d) の中からそれぞれ1つずつ選びなさい。
171. The heavy snowfall ( ) for the closing of school today.  
(a) excuses (b) causes (c) accounts (d) explains
  172. While I ( ) your point of view, I still think you are in the wrong.  
(a) deny (b) criticize (c) consider (d) respect
  173. The boy was ( ) injured in the accident.  
(a) dangerously (b) critically (c) greatly (d) heavily
  174. The government ( ) the plan to lower taxes.  
(a) mingled (b) rejected (c) deteriorated (d) decreased

175. A ( ) of wood hit him in the face.  
(a) chip (b) crumb (c) blade (d) puff

**APENDIX II**  
GRAMMAR TEST

Choose the better word in the blanket.

1. The laws of science (have/has) no exceptions.
2. I (saw/have seen) the statue when I visited the museum.
3. If he had known about it, he (might/might have) come earlier.
4. It (may/must) rain tomorrow, but I hope it will be sunny.
5. My parents (were discussing/were discussed) my future.
6. We (discussed/discussed about) the matter for five hours.
7. They let the children (leave/to leave) school.
8. I was (boring/bored) by the lecture.
9. I'm looking forward to (see/seeing) you again.
10. (What/How) foolish girls you have been !
11. A number of the miners (was/were) still on strike.
12. They will leave as soon as he (arrives/will arrive) .
13. He orders me about as if I (am/were) his wife.
14. Sheila (must/may) work harder : if she doesn't, she'll fail her exams.
15. A lot of money (has wasted/has been wasted) by the club.
16. He (married/married to) a girl from the noble family.
17. All I did was (report/to report) the accident.
18. (Seen/Seeing) from a distance, it looked like a human face.
19. This pencil needs (sharpening/to sharpen) .
20. He asked me (if/when) I knew his father.
21. Either the two witnesses or the judge (was/were) wrong.
22. He showed me the camera which he (bought/had bought) the day before.
23. It's time we (are/were) leaving.
24. She lives in Paris but she speaks with a German accent. She (mustn't/can't) be French.
25. I (surprised/was surprised) at the news.
26. He got (home/to home) late last night.
27. I hope (seeing/to see) you soon.
28. What is the language (speaking/spoken) in South America ?
29. I never see you without (think/thinking) of my brother.
30. This chair looks (comfortable/comfortably) to sit.

APENDIX III

READING COMPREHENSION TEST

次の文の空所に、適切な英語の単語を一語入れなさい。

I. The Birth of Jesus

As Judaea was part of the Romann Empire. Mary and Joseph had to obey the laws of Rome. "Every man and woman who lives in our lands must pay tax," went the proclamation of the Emperor Augustus. "A new register will be drawn up, on which the name of every citizen is written. The Emperor wants a complete record of his people."

When King Herod heard this command, (1. ) ordered everyone in his kingdom to (2. ) to the place of their birth (3. ) be put on the register. Because (4. ) belonged to the family of David, he (5. ) to leave Nazareth where he lived (6. ) worked and go back to Bethlehem, (7. ) lay some 70 miles to the south.

(8. ) his wife. Mary had to go, (9. ) , although her child, was expected any (10. ). Together they set off on their mule (11. ) the hills to Bethlehem. In the (12. ) they arrived in the city, and (13. ) to find somewhere to stay. But (14. ) was poor and the streets were (15. ) with people.

They walked on until (16. ) came to an inn. Lamps burned (17. ) every window and the rooms inside (18. ) full of people eating and drinking. Joseph (19. ) on the door. The landlord opened (20. ) .

"No room at the inn!" he (21. ) out, and waved them away. As (22. ) moved on, the light from the (23. ) fell on Mary.

"You can sleep (24. ) the stable, if you like. The (25. ) there is fresh," the landlord called (26. ) her. So Mary made a bed (27. ) herself in the sweet-smelling hay. And (28. ) , with the ox and the donkey (29. ) her, she gave birth to her (30. ) . She wrapped him up well and (31. ) him in a manger filled with (32. ) . She called him Jesus, which (33. ) "God Saves."

That night the words of the prophet Isaiah came true.

III. The Post Office

A : Foreign student      B : Her American friend

A foreign student wants to buy some stamps and also send some money home.

A : Where's the nearest post office ?

B : It's three (1. ) away, on Main Street. Do you need (2. ) stamps ?

A : Yes, I need some air mail (3. ) and I want to mail some letters.

B : (4. ) don't have to go to the post (5. ) for that. You can buy stamps in (6. ) stamp machine at the drugstore, and you (7. ) mail your letters in the mail box.

A : (8. ) I think I'd better go to the (9. ) office. I've got six letters that I (10. ) to mail to my country. Besides, I (11. ) air mail stamps.

B : You can send air (12. ) letters with any kind of stamps, so (13. )

as you put the right amount on. (14. ) the way, do you know about air (15. ) ?

A : No, what are air letters ?

B : An air (16. ) is a special sheet of paper with (17. ) stamp printed on it. And you can (18. ) them anywhere in the world.

A : That's a (19. ) thing to know about. I'll get some (20. ) afternoon. By the way. I want to (21. ) some money to a friend of mine. (22. ) it safe to send cash in a (23. ) ?

B : It's not really a good idea. Do (24. ) have a bank account ?

A : No, not yet.

B : (25. ) you can't send a check. If I were you. I'd get a money order at the post office. That's the safest way to send money. Your friend can cash it like a check.

A : That sounds like a good idea.

I. 1) 次の文の各 ( ) の中から,それぞれ適切な語を1つ選び丸で囲みなさい。

Recently, I made a trip to San Francisco. I hadn't been particularly interested in going there, (1.and, but, so, in fact) my friend convinced me that it would be fun. We saw some famous tourist attractions. (2.For example, Later, And, Then), we saw the Golden Gate Bridge, Chinatown, and Ghirardelli Square. (3.In short, Then, Therefore, As a result), we went to the island of Alcatraz. One evening, we went to see a play. (4.However, In addition, To conclude, For instance), San Francisco has some good restaurants. (5.Nevertheless, Next, On the whole, So,) we enjoyed eating out. We also did some shopping, (6.although, so, and, of course) we didn't buy anything. (7.But, First, In fact, In summary) we had a really good time. (8.Even though, Consequently, On the whole, Naturally,) it was one of the nicest vacations I've ever had.

#### ADENDIX IV

#### PRONUNCIATION & SPEAKING

One day, Yoshiko and her mother were on their way downtown to see a movie. While they were waiting for a bus, they saw a foreigner. When the bus came, he asked the bus driver something in English. But the driver couldn't understand him. Her mother told Yoshiko to speak to the foreigner in English. He understood her English. Then Yoshiko Discovered that he was going to see the same movie. She told him that she would guide him to the movie theater. He was very glad to hear that. She was also very happy to speak English with the foreigner.

#### Questions

No. 1 Where were Yoshiko and her mother on their way to ?

No. 2 What were they waiting for ?

No. 3 Could the driver understand the foreigner ?

No. 4 What did Yoshiko discover when she spoke to the foreigner ?

No. 5 Why was she happy ?