IJCLA VOL. 3, NO. 2, JUL-DEC 2012, PP. 77–92
RECEIVED 30/10/11 ACCEPTED 09/12/11 FINAL 08/06/12

Corpus Materials for Constructing Learner Corpus Compiling Speaking, Writing, Listening, and Reading Data

KATSUNORI KOTANI,[1] TAKEHIKO YOSHIMI,[2]
HIROAKI NANJO,[2] AND HITOSHI ISAHARA[3]

[1] *Kansai Gaidai University, Japan*
[2] *Ryukoku University, Japan*
[3] *Toyohashi University of Technology, Japan*

ABSTRACT

This paper presents the corpus material of a learner corpus called the I-Learner corpus consisting of text and sounds that reflect the proficiency of learners of English as a foreign language with respect to speaking, writing, reading, and listening, along with the types and quantity of the corpus materials. In constructing a learner corpus, a prerequisite is to prepare corpus materials that properly reveal learners' second language ability. Most conventional learner corpora use corpus materials taken from linguistic exercises such as essay writing and speaking exercises. The I-Learner corpus is the first corpus that collects the four-modality data, and the focus of this study is the selection of its material.

KEYWORDS: *Learner corpus, corpus materials, four-modality data*

1. INTRODUCTION

Learner corpora, which are defined as a collection of texts produced by learners of a second or foreign language [1], have contributed to the advancement of research on second language learning and teaching by providing text and sounds to analyze which linguistic items, such as vocabularies and grammars, learners adequately or inadequately use.

Some learner corpora [2, 3] are annotated with information tags on errors that learners made, thus making it possible to directly analyze learners' errors and/or compare the errors across learners of different proficiency levels. Learner corpora can also be used as a language resource in constructing computer-based language learning or teaching systems by machine learning algorithms [4].

The construction of a learner corpus consists of three steps: design, data collection, and analysis of collected data. The design step determines variables of a corpus. For example, the focus could be on language-related variables, task-related variables, and/or learner-related variables [5]. In the data collection step, raw text, sound, and information to be annotated with the text, such as learner information and error information, are collected. In the analysis of collected data step, basic analyses are performed, such as descriptive statistics analysis or qualitative analysis, to confirm the validity of the collected data.

Most learner corpora consist of text and sounds that reflect learners' proficiency in either writing [6] or speaking [2], but some include text that reflects learners' proficiency in the multiple modalities of speaking, writing, reading, and listening [7, 8, 9]. Wen et al. [7] constructed a learner corpus consisting of text that reflects learners' proficiency in speaking and writing. The speaking data included sounds and text transcribed from what learners had verbalized in speaking exercises, and the writing data included text from learners' essays. Meurers et al. [8] constructed a learner corpus consisting of text that reflects learners' reading and writing proficiency. The data included text written by learners as answers for comprehension questions in reading exercises. Kotani et al. [9] constructed a learner corpus, called the I(ntegrated)-Learner corpus, consisting of text and sounds that reflect learners' speaking (with a focus on pronunciation), writing, reading, and listening proficiency. According to them [9], one of the goals of this corpus is to provide a language resource for the analysis of learners' language use based on the four modalities because there is no other learner corpus that currently does so.

In constructing any learner corpus, the basic prerequisite is to select corpus materials that properly reveal learners' second language ability. Therefore, previous corpora have used materials taken from linguistic exercises such as essay writing [6, 7] and language tests [2, 7, 8, 9]. However, we feel that the selection of the corpus material of the I-Learner corpus [9] should be described in more detail because it is the first corpus that collects the four modality data. Therefore, in this paper

we discuss its design at length and also describe the types and quantity of the corpus materials.

2. I-LEARNER CORPUS

2.1. *Fundamental Design*

The I-learner corpus [9] was constructed on basis of the following design criteria: modality, context, technicality, data to be collected, learner, and task. In this subsection, we describe the modality, context, technicality, and data to be collected; the other criteria are described in the following subsections.

The modality consists of speaking, writing, listening, and reading. The context is the expository language used in daily-life contexts. The technicality is kept as low as possible in order to focus on linguistic proficiency. The data to be collected consist of language production data, language comprehension data, and mental language processing data.

The data to be collected are summarized in Table 1. The language production data, which show what the learners have produced, include both the sound of speaking and written sentences. The language comprehension data include the comprehension rate, which shows how well the learners comprehend the content of a text. The mental language processing data, which show how learners produced or understood sentences and/or sounds , include the speaking time, the writing time, the reading time, and the subjective judgment score, which is obtained by using a psychological data collection method [10] and shows what the learners thought as they were using English. The subjective judgment score of speaking on a five-point scale represents the difficulty of a sentence for the learner who pronounced that sentence. The subjective judgment score of writing on a five-point scale represents the comprehensibility of an English sentence written by a learner. The subjective judgment scores of listening and reading on a five-point scale represent the comprehensibility of a sentence for a learner who listened to or read the sentence.

Table 1. Data to be collected

	Language production data	Language comprehension data	Mental language processing data
Speaking	Sound	—	Speaking time Subjective judgment score
Writing	Sentence	—	Writing time Subjective judgment score
Listening	—	Comprehension rate	Subjective judgment score
Reading	—	Comprehension rate	Reading time Subjective judgment score

2.2. *Learners*

Learners of English as a foreign language were recruited, with candidates submitting their scores of the Test of English for International Communication (TOEIC) taken within a year of the start of the data collection. Ninety learners were accepted so as to obtain the same number of learners in each of the three proficiency levels: beginner (N = 30, TOEIC score of 280–495), intermediate (N = 30, TOEIC score of 500–725), and advanced (N = 30, TOEIC score of 730–985). The learners' first language was Japanese, and their education level was a university degree or higher, meaning that all had at least 36 months learning experience.

2.3. *Tasks of Data Collection*

The learners completed tasks (language tests of the four modalities) in the following order: listening, reading, speaking, and writing. For all tasks, they used a data collecting tool that displayed a sentence on a computer screen. This tool kept track of time when a learner verbalized, wrote, and read each sentence. It provided comprehension questions and saved answers for the listening and reading tasks. In the writing tasks, it displayed pictures and questions as well as blank spaces in which to write sentences. It kept a subjective judgment score during all the tasks.

In the listening tasks, the learners listened to four news articles that were read aloud by native speakers of English. They judged the difficulty of comprehending a sentence after listening to it. When they

finished listening to a news article, they answered five comprehension questions.

In the reading task, the learners silently read four news articles (which were different from the ones used in the listening task). They judged the difficulty of comprehending a sentence after reading it. When they finished reading a news article, they answered five comprehension questions. The use of a dictionary was prohibited, and the learners were allowed to read a sentence only once.

In the speaking task, the learners verbalized sentences from the four news articles that were used in the reading task. The same news articles were used so that the learners could grasp the content before the task began, thus enabling them to focus on pronunciation. They judged the difficulty of speaking a sentence after verbalizing it. There were no comprehension questions, unlike in the listening and reading tasks, because the focus was entirely on pronunciation, not comprehension.

In the writing task, the learners first described four pictures that comprised a series of events. They were assigned to write at least five sentences per picture. Next, they were provided with 20 questions, which they then answered. Here, they were assigned to write at least one sentence per answer. They judged the comprehensibility of a sentence after writing it. The use of a dictionary was prohibited, and the learners were not permitted to rewrite a sentence after they had moved on to another.

2.4. *Collected Data*

There were 90 learners who listened to 80 sentences from 4 news articles and answered 5 comprehension questions for each news article. Therefore, the listening data consisted of 7200 sentences annotated with a subjective judgment score and 360 examples of comprehension rate.

The reading data consisted of 7200 sentences annotated with the reading time and the subjective judgment score and 360 examples of comprehension rate. The total reading time was approximately 25.5 hours.

The speaking data consisted of 7200 sentences annotated with the speaking time and the subjective judgment score. The total speaking time was approximately 28.9 hours.

The 90 learners were asked to write at least 40 sentences for the writing task, so the writing data consisted of at least 3,600 sentences annotated with the writing time and the subjective judgment score. The

total writing time for the picture description was approximately 28.4 hours and that for answering questions was 30.2 hours.

3. MATERIALS OF I-LEARNER CORPUS

The materials used in the I-Learner corpus [9] were selected on basis of the design criteria (modality, context, technicality, data to be collected, learner, and task) described in Section 2.

3.1. *Material Design*

In compiling the learners' language data, we determined the design of corpus materials to emphasize the contrast between success and failure in that data. We designed the corpus materials to include three types of linguistic properties that enhance the contrast: the syntactic property of sentence length, semantic property of question type, and discourse property of information structure.

The speaking, listening, and reading materials were designed to include different syntactic difficulties and semantic difficulties. We used sentence length as an index of syntactic difficulties. Sentence length leads to difficulty in comprehending or processing linguistic objects, as previous research on readability [11] has shown. Thus, the news articles in the speaking, listening, and reading materials should contain different sentence lengths.

We used the type of question, such as true questions, false questions, and content questions, as an index of semantic difficulties. The effect of the type of question on the learners' language data should be examined in future work, but we expect that the question types cause the following differences in semantic difficulty. Content questions should be more difficult to answer than true questions and false questions because answers cannot be determined in a binary way (true or false). The language learners have to recognize what the article is about to answer content questions. In contrast, answers to true questions and false questions can be determined in a binary way. In addition, false questions should be more difficult than true questions to answer because deciding the correct answer to false questions, which needs negative evidence, requires more logical thinking than finding positive evidence.

The writing materials were designed to include different discourse difficulties and semantic difficulties. We used the discourse direction and the number of people in a picture [12] as an index of discourse difficulties. The effect of the discourse difficulties on the learners' language data should be examined in future work, but we expect that the discourse direction and the number of people in a picture cause the following difference in discourse difficulty. When describing these pictures, the learners have to represent the situation following the discourse direction on the basis of a proper information structure [13]. That is, when a new person appears, the person should be treated as new information. However, this person should be treated as old information in the subsequent picture. Thus, multiple pictures in the writing materials should represent a series of events, and different combinations of people should appear in each picture.

We used the type of question, such as polar or wh-interrogatives, as an index of semantic difficulties. The effect of the type of question on the learners' language data should be examined in future work, but we expect that the question types cause the following difference in semantic difficulty. Questions asking for descriptive comments should be the most difficult for which to write answers. The second most difficult should be wh-interrogative-type questions, and the least difficult should be polar-interrogative-type questions. Thus, questions in the writing materials should include these three types of questions.

3.2. *Speaking, Listening, and Reading Materials*

The speaking, listening, and reading materials of the I-Learner corpus were compiled from news articles taken from the Voice of America (VOA) site (http://www.voanews.com). The articles were chosen in two steps. In the first step, special sections for English learners and editorial sections were chosen from the various ones available on VOA. The articles in the former should be easier than those in the latter. This is because articles in special sections for English learners in VOA are written in short, simple sentences that contain only a core vocabulary of 1,500 words and no idiomatic expressions, according to VOA, while articles in editorial sections are written for native English speakers in sentences that have no restrictions. In the second step, articles were chosen according to conditions on the article size (number of words in an article) being approximately 350 words (within plus or minus 5%) and on the number of sentences in an article being 25 sentences for

easy articles and 15 sentences for difficult articles. These conditions excluded the possibilities that easy articles contained more long sentences and that difficult articles contain more short sentences.

The same articles are used when compiling the speaking and reading data. First, the learners silently read four articles (two easy and two difficult ones), and then they read aloud those same articles. The first reading enables the learners to grasp the content of the articles. Thus, when reading aloud, they can focus on the pronunciation. Examples of an easy and a difficult article, respectively, are shown in Appendices 1 and 2. When reading an article silently or aloud, the learners see this article on a computer screen sentence by sentence.

The listening data are also compiled using four articles (two easy and two difficult ones). These articles were taken from the same sections of the VOA site as those used in the speaking and reading tasks. In addition, these articles met the conditions for the article size and number of sentences in an article. In the listening task, the learners listen to VOA reporters.

The linguistic properties of the articles used in the speaking and reading tasks are shown in Table 2, and the properties of the articles in the listening task are shown in Table 3. These tables provide the difficulty of the article (Difficulty: Easy or Difficult), the title of the article (Title), the number of words in an article (W), the number of words in the shortest sentence (Min), the number of words in the longest sentence (Max), the average number of words in the sentences (Mean), and the standard deviation (SD).

A one-way analysis of variance (ANOVA) was conducted to examine whether the sentence length (number of words per sentence), as an index for syntactic difficulties, differed between the easy and difficult articles. The article difficulty was determined based on the type of sections: special sections for English learners or editorial sections for native English speakers. There was a significant difference in the sentence length at the $p<.01$ level [$F(3, 76)=14.16$] in the articles for the speaking and reading tasks. Post-hoc comparisons using the Tukey honestly significant difference (HSD) test indicated that the mean values of the sentence lengths were significantly different between all the pairs of easy articles (E1, E2) and difficult articles (D1, D2). However, there was no significant difference between E1 and E2, or between D1 and D2.

Table 2. Properties of speaking and reading materials

Article ID	E1	E2	D1	D2
Difficulty	Easy	Easy	Difficult	Difficult
Title	Recruiters Help US Colleges Find Foreign Students	Book Predicts Jump in High School Courses Online	U.S. Designates Al-Quso Terrorist	Ending Impunity In the Congo
W	337	356	359	348
Min	7	5	12	11
Max	23	22	37	42
Mean	13.5	14.2	23.9	23.2
SD	4.6	4.2	7.7	10.1

Table 3. Properties of listening materials

Article ID	E3	E4	D3	D4
Difficulty	Easy	Easy	Difficult	Difficult
Title	Studying in the US: A Lesson in Personal Finance, Part 2	Studying in the US: Grading Grades	Educating Marginalized Children	Outreach To Muslims
W	358	341	357	353
Min	5	6	8	10
Max	22	20	39	38
Mean	14.3	13.6	23.8	23.5
SD	4.8	3.7	8.9	7.4

There was also a significant difference in the sentence length at the $p < 0.01$ level [$F(3, 76) = 16.22$] in the articles for the listening task. Post-hoc comparisons using the Tukey HSD test indicated that the mean values of the sentence lengths were significantly different between all the pairs of easy articles (E3, E4) and difficult articles (D3, D4). However, there was no significant difference between E3 and E4, or between D3 and D4. Taken together, these results show that the easy articles contain shorter sentences than the difficult articles.

The listening and reading materials included questions created by the author of this paper following question formats [14]. The questions

are categorized into three types: a question asking what is true, e.g., "Which of the following is mentioned?" (true question); what is false, e.g., "Which of the following is NOT mentioned?" (false question); and what the content is about, e.g., "According to the passage, why or how…?" (content question). Each article has two true questions, two false questions, and one content question. Appendix 3 illustrates the questions for the easy article shown in Appendix 1. The questions are multiple choices with four answer choices.

3.3. *Writing Materials*

In the picture description task, the learners describe a series of events. The events are represented in a series of four pictures (Appendix 4), and thus this material represents the discourse direction. Four people appear in these pictures. In picture A, a woman and a man appear. In picture B, a different man appears with the woman and man who appeared in picture A, for a total of three people. In picture C, only the two men appear. In picture D, a different woman appears with the other three people.

Given the discourse difficulties of the order of pictures and the number of people, describing picture D should be most difficult. The second-most difficult picture should be picture B or C. If the order of pictures contributes more to the difficulty of describing pictures, the difficulty of picture C would be greater than that of picture B. In contrast, if the number of people has a greater effect on the difficulty of describing pictures, picture B would be more difficult than picture C.

In the question answering tasks, the learners answer questions about their own learning profiles [15] and on their computer literacy [16] (Appendix 5). The sentences from 1 to 15 ask about the learners' learning profiles, and those from 16 to 20 ask about their computer literacy. Of these sentences, 13 are wh-interrogative-type and 5 are polar-interrogative-type questions. The remaining two sentences are not interrogatives; instead, they ask for descriptive comments.

4. CONCLUSION

The present paper introduced the corpus materials of the I-Learner corpus, which collected learners' language data for the four modalities of speaking, writing, listening, and reading. These materials were

designed to include different linguistic difficulties. The writing materials included different semantic difficulties and discourse difficulties: the type of question, the discourse direction, and the number of people in a situation. The speaking, listening, and reading materials included different semantic difficulties and syntactic difficulties: the type of question and the sentence length.

We further noted the expected effects of these linguistic difficulties on the learners' language data. However, we have not examined whether these effects appear in that data. This examination will provide fundamental information for assessing the validity of the corpus for future studies. Thus, one remaining issue is to examine whether the corpus materials actually emphasize the contrast between success and failure in learners' language data after compiling the relevant data.

REFERENCES

1. Tono, Y.: Integrating Learner Corpus Analysis into a Probabilistic Model of Second Language Acquisition. In P. Baker (ed.) Contemporary Corpus Linguistics. Continuum International Publishing Group, London, pp. 184–203 (2009).
2. Izumi, E., Uchimoto, K., Isahara, H. (eds.): Nihonjin 1200 Nin no Eigo Spiking Koopasu [A Speaking Corpus of 1200 Japanese Learners of English]. ALC Press, Tokyo, Japan (2004).
3. Gammon, M.: High-order Sequence Modeling for Language Learner Error Detection. Proceedings of the 6th Workshop on Innovative Use of NLP for Building Educational Applications, pp. 180—189, (2011).
4. Kotani, K., Yoshimi, T., Kutsumi, T., Sata, I., Isahara, H.: EFL Learner Reading Time Model for Evaluating Reading Proficiency. CICLing 2008, pp. 655–664 (2008).
5. Tono, Y.: Learner Corpora: Design, Development and Applications. Paper presented at the Corpus Linguistics 2003 Conference (CL 2003), (2003).
6. Granger, S., Dagneaux, E., Meunier, F., Paquot, M.: International Corpus of Learner English, version 2. Presses Universitaires de Louvain, Louvain-la-Neuve, Belgium, (2009).
7. Wen, Q., Liang, M., Yan, X.: Spoken and Written Corpus of Chinese Learners (SWECCL) 2.0. Foreign Language Teaching and Research Press, Beijing, China, (2008).
8. Meurers, D., Ott, N., Ziai, R.: Compiling a Task-based Corpus for the Analysis of Learner Language in Context. In Sam Featherston and Britta Stolterfoht, editors, Proceedings of Linguistic Evidence 2010, pp. 214–217, (2010).

9. Kotani, K., Yoshimi, T.: A Scoring Method for Second Language Writing based on Word Alignment. Proceedings of Pacific Association for Computational Linguistics (PACLING) 2011, (2011).
10. Lewis, C. H.: "Thinking Aloud" Method in Cognitive Interface Design. Technical Report IBM RC-9265, (1982).
11. Kate, R. J., Luo, X., Patwardhan, S., ranz, M., Florian, R., Mooney,R. J., Roukos, S., Welty, C.: Learning to Predict Readability Using Diverse Linguistic Features. Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010), pp. 546–554, (2010).
12. McCarthy, M.: Discourse Analysis for Language Teachers. Cambridge University Press, Cambridge, (1991).
13. Prince, E. F.: Toward a Taxonomy of Given-new Information. In Peter Cole, editor, Radical Pragmatics, Academic Press, New York, pp. 223–255, (1981).
14. Nation, P., Malarcher, C.: Reading for Speed and Fluency. Compass Publishing, Seoul, Korea, (2007).
15. Ehrman, M. E.: Understanding Second Language Learning Difficulties. SAGE Publications, London, (1996).
16. Eignor, D., Taylor, C., Kirsch, I., Jamieson, J.: Development of a Scale for Assessing the Level of Computer Familiarity of TOEFL Examinees. Research Reports RR98-7, Educational Testing Service, Princeton, New Jersey, (1998).

APPENDICES

Appendix 1. Easy Article in Speaking and Reading Tasks

01: College prices in the United States have been rising faster than other prices for thirty years or more.
02: Recently many of the nation's top colleges have agreed to increase their financial aid.
03: But one group often has to pay the full price for college: foreign students.
04: This may help explain why colleges are making greater efforts to recruit them.
05: Large universities are likely to use their own representatives.
06: But smaller schools may work with independent recruiters.
07: An example is Albright College in Reading, Pennsylvania.
08: It has about one hundred foreign students, mostly from Asia.
09: It offers foreign students a savings of one-fifth off its published price if they apply through Study Group Holdings.
10: This placement company operates the Web site go-study.com.
11: Albright's international student counselor, Nicole Christie, says the company is paid from the money that the students pay the college.

12: Study Group looks for qualified students and rates their English skills before they apply.
13: But foreign students themselves often pay recruiters.
14: The recruiters help them write applications, get recommendation letters and prepare for admissions tests.
15: And they might help students prepare for getting a visa to study in the United States.
16: Recruiters can also work for both students and colleges.
17: Some education officials call this a conflict of interest.
18: They wonder how recruiters can find a school that is truly right for a student when certain colleges are paying them.
19: Officials also warn that like any other business, there is a risk of dishonesty.
20: Recruiters say they provide a useful service that is legal in the United States.
21: They say the colleges they work for are accredited and provide a good education but may not be widely known.
22: Recruiting of foreign students has been the subject of recent stories in the Chronicle of Higher Education and in the New York Times.
23: We are interested in hearing about experiences with college recruiters.
24: Send us your comments and we may use them in a future report.
25: Write to special@voanews.com and please include your name and country.

Appendix 2. Difficult Article in Speaking and Reading Tasks

01: The United States and the United Nations have listed Al-Qaida in the Arabian Peninsula fugitive Fahd al-Quso as a Specially Designated Terrorist.
02: These actions will help stem the flow of finances to and inhibit the travel of this dangerous operative.
03: The designation of Fahd al-Quso highlights U.S. action against the threat posed to the United States by al-Qaida in the Arabian Peninsula, said U.S. Ambassador for Counterterrorism Daniel Benjamin.
04: The joint designation by the United States and the United Nations alerts the public that Fahd al-Quso is actively engaged in terrorism.
05: These actions, said Ambassador Benjamin, "expose and isolate individuals like al-Quso and result in denial of access to the global financial system."
06: Prior to the formation of al-Qaida in the Arabian Peninsula, or AQAP, al-Quso was associated with al-Qaida elements in Yemen and involved in the 2002 USS Cole bombing in the Port of Aden, which killed seventeen sailors.
07: He was jailed in Yemen in 2002 for his part in the attack.
08: Following al-Quso's release from prison in 2007, he joined al-Qaida in Yemen.
09: In November 2009, al-Quso was added to the list of the FBI's most wanted terrorists.
10: Al-Quso is connected to other designated AQAP senior leaders, including Anwar al-Awlaqi, Nasir al-Wahishi, and Said Ali al-Shiri, and acts as a cell leader in Yemen.

11: In May 2010, al-Quso appeared in an al-Qaida in the Arabian Peninsula video in which he threatened to attack the U.S. homeland, as well as U.S. embassies and naval vessels abroad.
12: The terrorist designation blocks all al-Quso's property interests subject to U.S. jurisdiction and prohibits U.S. citizens from engaging in transactions that benefit al-Quso.
13: In addition to the U.S. domestic action, the United Nations Sanctions Committee's listing will require all U.N member states to implement an assets freeze, a travel ban, and an arms embargo against al-Quso.
14: The actions taken against the AQAP operative demonstrate international resolve in eliminating its ability to execute violent attacks and to disrupt, dismantle, and defeat their networks.
15: This designation represents just one phase of the U.S. government's response to the threat posed by al-Qaida in the Arabian Peninsula.

Appendix 3. Comprehension Questions for Easy Article Shown in Appendix 1

1. Which of the following is mentioned?
   (a) College teams from around the world took part in a computer programming competition.
   (b) Second of two reports on the business of bringing together students and schools.
   (c) Wealthier countries agree to limit how aggressively they recruit from developing countries.
   (d) Placement companies may be paid by colleges or students -- or both, raising concerns about possible conflicts of interest.

2. Which of the following is mentioned?
   (a) Universities will make greater efforts to recruit foreign students.
   (b) Universities agreed to increase their financial aid for foreign students.
   (c) Universities operate the Web site go-study.com.
   (d) Universities are interested in hearing about experiences with college recruiters.

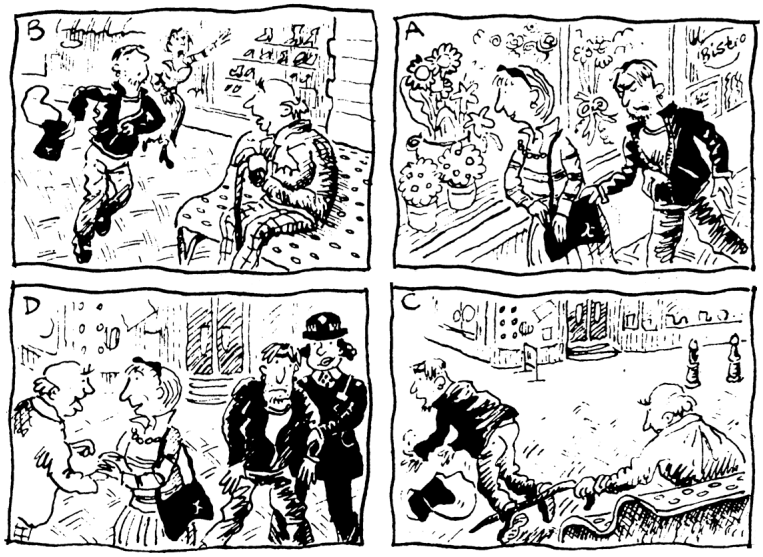
3. According to the passage, why do universities make efforts to recruit foreign students?
   (a) Because college prices have been rising.
   (b) Because universities work with independent recruiters.
   (c) Because foreign students have to pay the full price for college.
   (d) Because universities look for qualified students.

4. Which of the following is NOT mentioned?
   (a) A college offers foreign students a savings of one-fifth off its published price.
   (b) Recruiters help foreign students prepare for admissions tests.

(c) Recruiters work for both students and colleges.
(d) Large universities work with independent recruiters.

5. Which of the following is NOT mentioned?
   (a) Recruiters provide a useful service that is illegal in the United States.
   (b) Recruiters help foreign students prepare for getting a visa to study in the United States.
   (c) Some colleges providing a good education may not be widely known.
   (d) You can send them your comments.

Appendix 4. Pictures for Description



Appendix 5. Sentences for Question Answering

1. What were your favorite subjects?
2. What were your least favorite subjects?
3. What were your TOEIC scores (most recent)?
4. When did you last attend a class or take a course of any sort?
5. What was the class?
6. Which languages do you speak and read, and how well?
7. What language did you learn?
8. How did you learn the language?
9. How long did you learn the language?

10. Did you enjoy it?
11. Were you ever in contact with other languages while growing up? If yes, please describe briefly.
12. Did you find learning foreign languages easy?
13. Is there anything that might interfere with your learning and using another language?
14. Please add any additional comments about your past or anticipated language learning experience that might be helpful.
15. A variety of techniques may be used to help you learn foreign languages, by you and by your teachers. Please describe them.
16. How often is there a computer available for you to use at home?
17. How comfortable are you with using a computer?
18. How comfortable are you with using a computer to write a paper?
19. How many examinations/tests have you taken on a computer?
20. How often do you use a computer to send or receive e-mail?

KATSUNORI KOTANI
KANSAI GAIDAI UNIVERSITY,
16-1 NAKAMIYAHIGASHINO-CHO,
HIRAKATA, OSAKA, 573-1001, JAPAN
E-MAIL: <KKOTANI@KANSAIGAIDAI.AC.JP>

TAKEHIKO YOSHIMI
RYUKOKU UNIVERSITY,
1-5 YOKOYA SETA OE-CHO,
OTSU, SHIGA, 520-2194, JAPAN

HIROAKI NANJO
RYUKOKU UNIVERSITY,
1-5 YOKOYA SETA OE-CHO,
OTSU, SHIGA, 520-2194, JAPAN

HITOSHI ISAHARA
TOYOHASHI UNIVERSITY OF TECHNOLOGY,
1-1 HIBARIGAOKA, TEMPAKU,
1-2 TOYOHASHI, AICHI, 441-8580, JAPAN