Title	A Corpus of Writing, Pronunciation, Reading, and Listening by Learners of English as a Foreign Language
Author(s)	Kotani, Katsunori; Yoshimi, Takehiko; Nanjo, Hiroaki; Isahara, Hitoshi
Citation	English Language Teaching (ELT) (2016), 9(9): 139-155
Issue Date	2016-09
URL	http://hdl.handle.net/2433/229411
Right	Copyright for this article is retained by the author(s), with first publication rights granted to the journal.; This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).
Type	Journal Article
Textversion	publisher

A Corpus of Writing, Pronunciation, Reading, and Listening by Learners of English as a Foreign Language

Katsunori Kotani[1], Takehiko Yoshimi[2], Hiroaki Nanjo[3] & Hitoshi Isahara[4]

[1] School of Foreign Studies, Kansai Gaidai University, Osaka, Japan

[2] Faculty of Science and Technology, Ryukoku University, Shiga, Japan

[3] Academic Center for Computing and Media Studies, Kyoto University, Kyoto, Japan

[4] Information and Media Center, Toyohashi University of Technology, Aichi, Japan

Correspondence: Katsunori Kotani, School of Foreign Studies, Kansai Gaidai University, Hirakata, Osaka, 573-1001, Japan. Tel: 81-72-805-2801. E-mail: kkotani@kansaigaidai.ac.jp

Abstract

In order to develop effective teaching methods and computer-assisted language teaching systems for learners of English as a foreign language who need to study the basic linguistic competences for writing, pronunciation, reading, and listening, it is necessary to first investigate which vocabulary and grammar they have or have not yet learned. Identifying such vocabulary and grammar requires a learner corpus for analyzing the accuracy and fluency of learners' linguistic competences. However, it is difficult to use previous learner corpora for this purpose because they have not compiled all the types of linguistic data that we need. Therefore, this study aimed to solve this problem by designing and developing a new learner corpus that compiles linguistic data regarding the accuracy and fluency of the four basic linguistic competences of writing, pronunciation, reading, and listening. The reliability and validity of the learner corpus were partially confirmed, and practical application of the learner corpus is reported here as case studies.

Keywords: English as a foreign language, learner corpus, accuracy, fluency, writing, pronunciation, reading, listening

1. Introduction

The globalization of society has increased opportunities to use English even in circumstances of English as a foreign language (EFL). The changing of outside the class has altered the purpose for EFL teaching in Asian countries (Kirkpatrick, 2012; Yamada, 2015). Before this changing appeared, EFL teaching primarily aimed at linguistic competences of reading and writing, that is, written language competences, but didn't concentrate on spoken language competences such as linguistic competences of listening and speaking. However, present EFL teaching aims to develop both spoken and written language competences (Canale & Swain, 1983).

In addition, current EFL teaching further aims to grow EFL learners who use accurately English at the natural speed (fluently). Although the fluency in terms of speed is not crucial in written language as EFL learners re-consider the use of vocabulary and grammar, it is crucial in spoken language because they need to understand English spoken at natural speed, and speak English at natural speed without reconsideration of vocabulary and grammar.

The two aims mentioned above are to be solved by providing EFL learners with basic practices for the use of vocabulary and grammar in listening, speaking, reading, and writing that fit the proficiency of each of EFL learners. In addressing this issue, it is necessary to first investigate EFL learners' basic linguistic competences to identify unlearned vocabulary and grammar for each of the linguistic competences. Recently, such identification of learner's linguistic problems has been made with error analyses through a learner corpus, because corpus-based analysis examines learner's language use. After that, effective teaching methods and computer-assisted language teaching (CALT) systems to improve communicative linguistic competences for spoken and written English can be developed.

This study primarily investigated previous learner corpus research in order to find a learner corpus that enables

analyses for multiple linguistic competences for each learner from the viewpoints of accuracy and fluency. However, most of the previous learner corpora targeted linguistic competences of writing and speaking (Minematsu et al., 2003; Nicholls, 2003; Izumi et al., 2004; Granger et al., 2009; Yasuda et al., 2009; Gilquin, 2010), while a few corpora compiled listening data (Rytting et al. 2014, ten Bosch et al. 2015) or reading data (Meurers et al., 2010; Ott et al., 2012). In addition, all of the previous learner corpora enabled analyses for the accuracy, but not for the fluency. Hence, it was concluded that we need to compile a learner corpus for analyzing the accuracy and fluency of multiple linguistic competences.

Given the present situation of EFL teaching and research on learner corpora, the present study compiled a learner corpus with two purposes: (1) to compile data for close investigation of EFL learners' accuracy and fluency in writing, pronunciation, reading, and listening; and (2) to compile data for the development of basic modules for a CALT system that can choose basic practices appropriate to the target EFL learners' proficiency level.

The proposed learner corpus of this study will contribute to develop automatic evaluation methods of learners' language use in different modes as well as the language use from both accuracy and the fluency, which lead to a better computer-assisted language learning system. The learner corpus will also benefit to analyze the learning process from these perspectives, and the findings of the analyses will contribute to develop automatic evaluation methods of learning materials that suit to the proficiency.

2. Previous Learner Corpora

This section introduces the principal learner corpora developed in previous research according to the linguistic competences targeted. Table 1 summarizes the previous learner corpora.

Table 1. Specifications of previous learner corpora

Author(s)	Target competence	Target analysis
Nicholls (2003)	Writing	Accuracy
Granger et al. (2009)		
Yasuda et al. (2009)	Writing	Accuracy & Fluency
Minematsu et al. (2003)	Pronunciation	Accuracy
Izumi et al. (2004)	Speaking	Accuracy
Gilquin (2010)		
Ott et al. (2012)	Reading & Writing	Accuracy
Meurers et al. (2010)		
Rytting et al. (2014)	Listening	Accuracy
Ten Bosch et al. (2015)		

First, we have learner corpora that compiled linguistic data in writing (Nicholls, 2003; Granger et al., 2009; Yasuda et al., 2009, among others). Data contained in these corpora are sentences written by EFL learners and annotated linguistic information such as part-of-speech information, syntactic information, and information on linguistic errors in word choice and compliance with grammatical rules. Hence, these corpora are effective for investigating vocabulary and grammar that EFL learners have not yet learned or find difficult to use. These corpora can be used for error analysis in order to investigate which sentences are accurately or inaccurately used by EFL learners. Among these corpora, the corpus of Yasuda et al. (2009) can also be used to investigate which sentences EFL learners can use fluently because it contains data on EFL learners' level of confidence in the accuracy of a sentence.

Second, we have learner corpora that compiled linguistic data on speaking (Minematsu et al., 2003; Izumi et al., 2004; Gilquin, 2010, among others). These corpora can be further classified into two types: one type recorded the speech sound of EFL learners reading aloud English sentences that were assigned by the researchers (Minematsu et al., 2003), and the other recorded both speech sound and sentences that they generated in conversation (Izumi et al., 2004; Gilquin, 2010). The former can be used to investigate pronunciation errors, and the latter to investigate both pronunciation errors and errors regarding vocabulary and grammar when generating a sentence.

Third, we have learner corpora that compiled linguistic data on reading (Meurers et al., 2010; Ott et al., 2012), although the target learners were learners of German as a foreign language, not EFL learners. Data contained in these corpora are sentences from reading materials and answers for reading comprehension questions written by the learners. These corpora can be used to analyze errors in both reading comprehension and writing.

Lastly, we have learner corpora that compiled linguistic data on listening (Rytting et al., 2014; ten Bosch et al., 2015), although the target learners were learners of Arabic and Dutch as a foreign language, respectively. In these corpora, error tags of dictation were annotated on target words in each sentence. The former can be used to develop a listening error-correction tool, and the latter to investigate problems in listening comprehension and develop reference data for evaluating a computational model of human word recognition.

3. Corpus Data Compilation

3.1 Data to Be Compiled

Table 2 shows the language use data that satisfied these two purposes. Written sentences (writing), speech sounds (pronunciation), and comprehension rate (in reading and listening) were used to analyze the accuracy of language use (error analysis). Other data, including processing rates (in writing, pronunciation, and reading) and confidence judgments were used to analyze the fluency of language use (fluency analysis).

Table 2. Language use and corpus data

Language use	Data for error analysis	Data for fluency analysis	
Writing	Written sentence	Processing rate	Confidence judgment
Pronunciation	Speech sound	Processing rate	Confidence judgment
Reading	Comprehension rate	Processing rate	Confidence judgment
Listening	Comprehension rate		Confidence judgment

In Table 2, the comprehension rate in reading and listening refers to the percentage of correct answers to comprehension questions for written or spoken text that an EFL learner silently read or listened to. The processing rate in writing, pronunciation, and reading refers to the speed at which an EFL learner wrote, read aloud, and silently read a sentence in terms of the number of words processed per minute (WPM). The confidence judgment in writing refers to an EFL learner's score of their confidence in the accuracy of a sentence they generated as rated on a five-point Likert scale (1: confident, 2: somewhat confident, 3: average, 4: not very confident, or 5: not confident), and the confidence judgment in the other data refers to the EFL learner's score of the difficulty of sentence pronunciation or comprehension as rated on a five-point Likert scale (1: easy, 2: somewhat easy, 3: average, 4: somewhat difficult, or 5: difficult).

3.2 Language Use Data Collection Method

To collect language use data, the EFL learners were asked to perform four tasks in the order shown in Table 3. After the reading task, the EFL learners were given a 60-minute break.

Table 3. The four tasks used to collect language use data

Task	Materials	Allotted time
1. Listening	80 sentences (four news clips)	8 minutes per clip (32 minutes for the task)
2. Reading	80 sentences (four news articles)	13 minutes per article (52 minutes for the task)
3. Pronunciation	80 sentences (four news articles)	9 minutes per article (36 minutes for the task)
4. Writing	Four pictures and 20 comprehension questions	30 minutes per material (60 minutes for the task)

In the listening task, EFL learners listened to four audio news clips sentence-by-sentence only once, using a data-collecting tool with headphones. After listening to each sentence, they estimated their confidence in their comprehension of the sentence and selected their confidence judgment on the five-point Likert scale explained in Section 3.1. When finishing a clip, they answered five multiple-choice comprehension questions. Multiple choices for confidence judgments and comprehension questions were shown on the data-collecting tool's computer screen, which also recorded the EFL learners' choices.

The reading task proceeded similarly to the listening task. There were four news articles, which were read silently sentence-by-sentence, and the EFL learners gave their confidence judgments for their comprehension of each sentence on the same five-point scale as above, and answered comprehension questions.

The pronunciation task also proceeded similarly to the reading task except comprehension questions were absent and the task was performed at a different location. Comprehension questions were omitted because the articles were the same as those used in the reading task. EFL learners read sentences aloud using the data-collecting tool with a unidirectional electric-condenser microphone mounted on a stand placed in front of the mouth in a sound-attenuated recording booth (width: 1700mm, depth: 1900mm, and height: 2100mm). Speech sound data were recorded on a solid-state stereo with a sampling rate of 44.1 KHz and a quantization level of 16 bits.

In the writing task, EFL learners were asked to write sentences describing four pictures and to answer 20 questions about their own background and computer skills. The EFL learners were asked to write at least, five sentences for each picture, and at least one sentence for each background question. After writing each sentence, they determined their confidence judgments as above. The data-collecting tool's screen showed the pictures and questions, and recorded the sentences, writing time and choices for confidence judgments.

In these data collection tasks, EFL learners were asked to complete each task as fast as possible during the allotted time, and to stop working either when the task was completed or when the experimenter and data-collecting tool alerted them of the end of the allotted time. EFL learners were prohibited from using dictionaries or any other reference books, and the data-collecting tool did not allow EFL learners to return to review a sentence after moving on to another sentence in the listening, reading and pronunciation tasks, or to revise or modify a sentence in the writing task.

3.3 Learners

90 EFL learners who were enrolled either in undergraduate or graduate programs at the university (males: 48 and females: 42) took part in the data collection tasks and were paid money as a reward for their participation. EFL learners at university were chosen because although Japanese students learn English for 6 years in junior and senior high school, low linguistic competences among Japanese students has been recognized as a problem leading universities to offer remedial English classes. The mean age of EFL learners in the present study was 21.5 years [range 19-40 years, standard deviation (SD) 2.6]. The EFL learners were asked to submit valid Test of English for International Communication (TOEIC) scores, taken in the current or previous year. The learners confirmed that they had basic computer skills such as typing with a keyboard and controlling a mouse.

The EFL learners were classified into three groups based on their TOEIC scores, each of which consisted of 30 learners. The TOEIC scores in the beginner group ranged from 280 to 485, those in the intermediate group ranged from 490 to 725, and those in the advanced group ranged from 730 to 985. Table 4 shows the descriptive statistics for the EFL learners' TOEIC scores. Figure 1 shows the distribution of the EFL learners' TOEIC scores.

Table 4. EFL learners' TOEIC scores

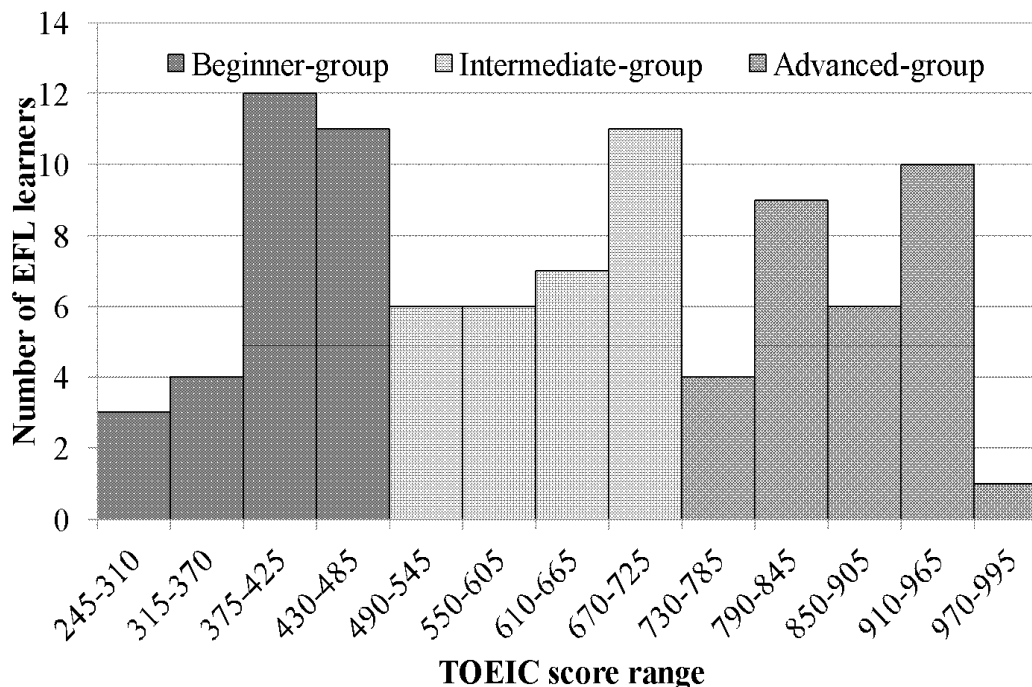|	*Mean (SD)*
All learners (n = 90)	633.8 (198.3)
Beginner group (n = 30)	404.7 (51.3)
Intermediate group (n = 30)	632.5 (69.8)
Advanced group (n = 30)	864.3 (69.0)

Figure 1. Distribution of TOEIC scores

Comparing the mean TOEIC score of the EFL learners in this study (633.8) with that of EFL learners in all of Japan (583.5) provided by Educational Testing Service showed that the participants in this study had a higher mean score. The difference in the mean scores likely arose from the fact that the EFL learners in this study were limited to undergraduate or graduate students, while the nationwide mean also included junior and senior high school students' scores. Thus, our learner corpus should be regarded as representative data for EFL learners at the university level.

3.4 Materials

The purpose of our learner corpus was to provide a base for researching EFL learners' grammatical competence with respect to sentence comprehension and generation. Therefore, completing each task required only basic grammatical competence, not advanced linguistic competence on rhetoric and discourse styles or non-linguistic knowledge such as background or cultural knowledge.

Materials for the listening, reading and pronunciation tasks were chosen by focusing on the use of grammatical competence. News articles were chosen for these tasks because the target EFL learners were undergraduate or graduate students who needed to learn vocabulary and grammar used in authentic materials such as news articles, and because news articles are often used as teaching materials.

The news articles were taken from the two main sections of the Voice of America (VOA) website (http://www.voanews.com): the special section prepared for English learners, and the editorial section prepared for native English speakers. Articles posted in the special section contain short, simple sentences consisting of the 1500 basic vocabulary of VOA, and avoid idiomatic expressions. In contrast, there are no restrictions on vocabulary or sentence construction in articles posted in the editorial section as long as they are appropriate as news articles for native English speakers.

The news articles and clips for the listening task were also taken from the special and editorial sections of the VOA website. In addition to the lexical and syntactic restrictions on special section news articles, news clips are phonetically restricted, and the speed of news clips posted in the special section is two-thirds slower than those in the editorial section, which are read aloud at a native English speaker's natural speed of approximately 250 syllables per minute (Robb &Gillon, 2007).

The same news articles were used as the materials for the reading and pronunciation tasks. This overlap was necessary to allow the EFL learners to focus on pronunciation, and to minimize the influence of the article

contents on learners' linguistic performance.

Tables 5 and 6 summarize the properties of the selected materials for the listening, reading and pronunciation tasks. The mean number of words in a sentence shows that sentences from the editorial sections are longer than those from the special section. This is also evident from the length of the shortest and longest sentences. Tables A and B (Appendix) provide examples of news articles for the reading task from the special section and editorial section, respectively.

Table 5. Properties of the materials selected for the listening task

Section	Title of news clip	Length of news clip				
		Total number of words	*Total number of sentences*	*Mean sentence length (number of words)* *(SD)*	*Shortest sentence length (number of words)*	*Longest sentence length (number of words)*
Special section	Studying in the US: A Lesson in Personal Finance, Part 2	358	25	14.32 (4.8)	5	22
Special section	Studying in the US: Grading Grades	341	25	13.6 (3.7)	6	20
Editorial section	Educating Marginalized Children	356	15	23.7 (8.9)	8	39
Editorial section	Outreach To Muslims	352	15	23.5 (7.6)	9	39

Table 6. Properties of the materials selected for the reading and pronunciation tasks

Section	Title of news article	Length of news article				
		Total number of words	*Total number of sentences*	*Mean sentence length (number of words)* *(SD)*	*Shortest sentence length (number of words)*	*Longest sentence length (number of words)*
Special section	Recruiters Help US Colleges Find Foreign Students	335	25	13.4 (4.7)	7	23
Special section	Book Predicts Jump in High School Courses Online	356	25	14.2 (4.2)	5	22
Editorial section	U.S. Designates Al-Quso Terrorist	350	15	23.3 (7.2)	12	37
Editorial section	Ending Impunity in the Congo	348	15	23.2 (10.1)	11	42

Each news article or clip contained five multiple-choice comprehension questions following the format of Nation and Malarcher (2007): two correct answer options; two incorrect answer options; and an alternative option. An example of the comprehension questions used is provided in Table C (Appendix).

Similar to the selection of materials for the other tasks, two types of materials were chosen in order for EFL

learners to complete the writing task by using basic grammatical competence, not advanced linguistic competence on rhetoric and discourse styles or non-linguistic knowledge such as background or cultural knowledge: pictures to describe and questions to answer. Four pictures representing the flow of an event on the street where four people appear were taken from a writing test (Hughes, 2003). These pictures were chosen because no special terminology was required to describe the event (Figure A, Appendix).

The EFL learners were asked 15 questions about their academic background, taken from Ehrman (1996), and five questions about their computer literacy, taken from Eignor et al. (1998) (Table D, Appendix). These questions were chosen because no special terminology was required to answer these questions, and the EFL learners could answer questions about their own background more easily than questions about other topics such as business or academic issues. The questions consisted of thirteen constituent interrogative questions, five polar interrogative questions, and two command questions.

3.5 Amount of Language Use Data

Table 7 summarizes the language use data collected in the writing, pronunciation, reading, and listening tasks, which satisfy the first two purposes of this study regarding the types of language use data to be compiled. The writing data collected consisted of more than 4,000 sentences (29,115 words); this is 1.1 times the minimum number of 3,600 sentences to be collected (90 EFL learners × 40 sentences each). On the other hand, the pronunciation, reading, and listening data lacked 7, 18, and 396 data pieces, respectively, from the expected 7,200 data pieces (90 EFL learners × 80 sentences each), respectively. The total length of the speech sound data collected was 28.9 hours.

Table 7. Language use data of the EFL learners

Language use	Data	Amount (pieces)
Writing	Sentences (number of words)	4,007 (29,115)
	Processing rate	4,007
	Confidence judgment	4,007
Pronunciation	Speech sounds (hours)	7,193 (28.9)
	Processing rate	7,193
	Confidence judgment	7,193
Reading	Comprehension rate	360
	Processing rate	7,182
	Confidence judgment	7,182
Listening	Comprehension rate	360
	Confidence judgment	6,804

4. Validation of Language Use Data

4.1 Reliability

We assessed the reliability of the language use data following a validation procedure for language tests (Brown 1996). Reliability was determined by investigating the extent to which language use data were consistent among EFL learners. This is because reliable data should yield similar results for EFL learners at the same proficiency level.

Reliability and internal consistency were analyzed in terms of Cronbach's alpha reliability coefficient (Cronbach 1970), which was mathematically defined using Equation 1, where α is a reliability coefficient, k is the number of items on language use data (the number of data items, $k = 9$), S_i^2 is the variance associated with item i, and S_T^2 is the variance associated with the sum of all k-item values. Cronbach's alpha reliability coefficient, which ranges from 0 (absence of reliability) to 1 (absolute reliability), is empirically judged as satisfactory if it is above 0.8.

$$\alpha = \frac{k}{k-1}\left(1 - \frac{1}{S_T^2}\sum_{i=1}^{k} S_i^2\right) \tag{1}$$

A high reliability coefficient ($\alpha = 0.89$) was observed. This reliability coefficient indicates that the language use data were 89% consistent, or reliable, with 11% random variance. This suggests a high reliability of the collected language use data.

4.2 Construct Validity

We assessed the validity of the language use data following a validation procedure for language tests (Brown 1996). Validity was analyzed by construct validity and criterion-related validity.

Construct validity was determined by investigating the degree to which the language use data could classify the EFL learners into a beginner, intermediate, or advanced group based on their level of English proficiency. Whether the language use data could classify the EFL learners into the three proficiency level groups was investigated with analysis of variance (ANOVA) with $p \leq 0.001$ representing statistical significance.

Tables 8, 9, and 10 summarize the means and SDs of the language use data for the three proficiency level groups. The mean values were calculated by dividing the sum of the data values for each task by the number of EFL learners per group ($n = 30$).

Table 8. Comprehension rate (%) of the three proficiency level groups

Task	Beginner group	Intermediate group	Advanced group
	Mean (SD)	*Mean (SD)*	*Mean (SD)*
Reading	47.3 (12.2)	52.0 (10.3)	65.8 (11.3)
Listening	43.3 (13.7)	49.8 (10.2)	67.0 (11.1)

Table 9. Confidence judgments of the three proficiency level groups

Task	Beginner group	Intermediate group	Advanced group
	Mean (SD)	*Mean (SD)*	*Mean (SD)*
Writing	2.07 (0.87)	2.02 (0.50)	1.36 (0.97)
Pronunciation	2.46 (0.59)	2.13 (0.46)	1.57 (0.79)
Reading	2.15 (0.70)	1.58 (0.47)	1.06 (0.61)
Listening	2.45 (0.64)	1.99 (0.53)	1.07 (0.85)

Table 10. Processing rate (WPM) of the three proficiency level groups

Task	Beginner group	Intermediate group	Advanced group
	Mean (SD)	*Mean (SD)*	*Mean (SD)*
Writing	7.71 (2.19)	9.44 (2.41)	14.12 (5.76)
Pronunciation	67.96 (10.75)	78.81 (12.72)	93.15 (10.99)
Reading	85.45 (28.27)	97.81 (28.12)	131.18 (41.32)

The one-way ANOVA showed statistically significant differences among the three proficiency level groups in every type of language use data, as shown in Table 11.

Table 11. Results of ANOVA

Data type		*F-value*	*p-value*
Comprehension rate	Reading	F (2, 87) = 21.72	$p \leq 0.0001$
	Listening	F (2, 87) = 32.46	$p \leq 0.0001$
	Writing	F (2, 87) = 7.41	$p \leq 0.0001$
Confidence judgment	Pronunciation	F (2, 87) = 15.67	$p \leq 0.0001$
	Reading	F (2, 87) = 24.72	$p \leq 0.0001$
	Listening	F (2, 87) = 32.46	$p \leq 0.0001$
	Writing	F (2, 87) = 22.56	$p \leq 0.0001$
Processing rate	Pronunciation	F (2, 87) = 36.11	$p \leq 0.0001$
	Reading	F (2, 87) = 15.27	$p \leq 0.0001$

We further examined the language use data using Tukey post-hoc comparisons among the three proficiency level groups. The results indicated that the advanced group showed significantly higher values than the intermediate and beginner groups in every type of language use data (Table 12). The beginner and intermediate groups showed statistically significant differences ($p \leq 0.01$) only for the confidence judgment in reading and the processing rate of pronunciation.

Table 12. Results of Tukey post-hoc comparisons

Data type		Result
Comprehension rate	Reading	Beginner vs. Intermediate: non-significant
		Beginner vs. Advanced: $p \leq 0.01$
		Intermediate vs. Advanced: $p \leq 0.01$
	Listening	Beginner vs. Intermediate: non-significant
		Beginner vs. Advanced: $p \leq 0.01$
		Intermediate vs. Advanced: $p \leq 0.01$
	Writing	Beginner vs. Intermediate: non-significant
		Beginner vs. Advanced: $p \leq 0.01$
		Intermediate vs. Advanced: $p \leq 0.01$
Confidence judgment	Pronunciation	Beginner vs. Intermediate: non-significant
		Beginner vs. Advanced: $p \leq 0.01$
		Intermediate vs. Advanced: $p \leq 0.01$
	Reading	Beginner vs. Intermediate: $p \leq 0.01$
		Beginner vs. Advanced: $p \leq 0.01$
		Intermediate vs. Advanced: $p \leq 0.01$
	Listening	Beginner vs. Intermediate: non-significant
		Beginner vs. Advanced: $p \leq 0.01$
		Intermediate vs. Advanced: $p \leq 0.01$
	Writing	Beginner vs. Intermediate: non-significant
		Beginner vs. Advanced: $p \leq 0.01$
		Intermediate vs. Advanced: $p \leq 0.01$
Processing rate	Pronunciation	Beginner vs. Intermediate: $p \leq 0.01$
		Beginner vs. Advanced: $p \leq 0.01$
		Intermediate vs. Advanced: p ≤ 0.01
	Reading	Beginner vs. Intermediate: non-significant
		Beginner vs. Advanced: $p \leq 0.01$
		Intermediate vs. Advanced: $p \leq 0.01$

These results suggested that the construct validity of the language use data is still open to doubt since statistically significant differences were not observed in most of the language use data between the beginner and intermediate groups. On the other hand, the results confirmed the construct validity between the beginner and advanced groups, and between the intermediate and advanced groups. Specifically, EFL learners in the advanced group could read and listen to English sentences more accurately than those in the beginner and intermediate groups, while accuracy in reading and listening comprehension was similar between the beginner and intermediate groups. In addition, the advanced group could read, listen to, write, and pronounce English sentences more fluently than the other two groups, and the intermediate group could read and pronounce English sentences more fluently than the beginner group. However, the writing fluency was similar between the beginner and intermediate groups.

4.3 Criterion-related Validity

Criterion-related validity was determined by comparing to what extent the language use data correlated with the scores of a well-established English test, TOEIC. According to the TOEIC technical manual (Chauncey Group International 1998), TOEIC scores correlate with those of other tests, such as the Comprehensive Adult Student Assessment System, the Test of English as a Foreign Language, and the Canadian Language Benchmark Assessment, with correlation coefficients ranging from 0.73 to 0.87. In the present study, we set the lowest correlation coefficient ($r=0.73$) as the target correlation coefficient between the language use data and the EFL learners' TOEIC scores.

The language use data demonstrated marginal criterion-related validity. A statistically significant correlation was observed between the language use data and TOEIC scores; absolute Spearman rank-order correlation coefficients ranged from 0.36 to 0.73 ($p < 0.01$), as shown in Table 13.

Table 13. Absolute correlation coefficients

Data type		r
Comprehension rate	Reading	0.61
	Listening	0.68
	Writing	0.36
Confidence judgment	Pronunciation	0.56
	Reading	0.63
	Listening	0.64
	Writing	0.66
Processing rate	Pronunciation	0.73
	Reading	0.61

These results suggest moderate criterion-related validity of the language use data excluding data for confidence judgments in writing (r = 0.36). This low correlation is likely due to the fact that the EFL learners in the present study could choose the sentences they wanted to write, unlike in the pronunciation, reading, and listening tasks in which they had to use sentences provided by the researchers. Hence, the confidence judgments in writing should be independent from the proficiency of the EFL learners. On the other hand, the processing rate in writing showed a high correlation, which suggests that the processing rate data should be used in order to analyze learners' writing competences.

5. Usability Evaluation of Language Use Data

This section discusses the usability of the language use data from the viewpoint of providing a research base for the development of basic modules of a CALT system. The language use data collected in the present study (Table 2) can serve as training data for the statistical development of various basic modules of a CALT system. Among the various possibilities, we actually developed two basic modules: statistical methods for measuring the readability or listenability of a sentence (Kotani et al., 2012, 2014), which are reported in this section as case studies.

In English teaching, it is necessary to choose materials with a level of readability or listenability that is

appropriate to the target EFL learners' proficiency since inappropriate materials will negatively affect teaching efforts and decrease motivation among learners. The readability or listenability of authentic materials is usually unknown, unlike materials in textbooks prepared specifically for EFL learners.

Using previously developed automatic readability or listenability measuring methods for authentic materials is problematic because EFL learners' proficiency is not taken into account. Unlike native speakers, EFL learners have different proficiencies; hence, it is necessary to take individual differences of EFL learners' proficiency into account when measuring readability or listenability.

To fix this problem, Kotani et al. (2012, 2014) developed methods to measure the readability or listenability of a sentence according to EFL learners' proficiency using multiple regression analysis. To account for individual differences, Kotani et al. (2012, 2014) considered scores on the difficulty of sentence comprehension among EFL learners to be dependent variables in the multiple regression analysis, and EFL learners' proficiency to be one of the independent variables (the other independent variables were properties representing the linguistic complexity of a sentence). The scores judged by the EFL learners reflected the individual differences in proficiency because they were subjectively judged by the EFL learners themselves, as discussed in Section 3.2, while the EFL learners' proficiency was represented in terms of TOEIC scores, as discussed in Section 3.3.

6. Conclusion

Motivated by the present situation of EFL teaching and previous research on learner corpora, the present study compiled a learner corpus that could (1) demonstrate the accuracy and fluency of EFL learners' language use to comprehend a sentence when reading and listening, and to generate a sentence in writing and speaking (pronunciation); and (2) serve as a language resource for the statistical development of various basic modules of a CALT system.

What this study implies is that learner corpus study should examine different linguistic skills from different perspectives. Since EFL learners are on the learning process, their linguistic skills have not yet been stable. Thus, an EFL learner may succeed in reading comprehension of a sentence, but fail in fluently reading aloud of the sentence. As shown in Table 12, the beginner and intermediate groups read aloud English sentences with the different processing rates, but no statistically significant difference was observed in their processing rates of writing and reading. If the learner corpus study compiles the language use data in different linguistic skills from different perspectives, the contribution of learner corpus data to the development of CALT system will be enhanced more.

A remaining problem of our learner corpus is that it needs to be extended quantitatively and qualitatively. These improvements are necessary in order to gain more detailed information with respect to EFL learners and their language use. Under our learner corpus, the language use data targeted the four types of language use of the three proficiency levels of EFL learners; in the future, corpus data should further be compiled for analyzing different language uses such as essay writing, speaking, narrative reading, and listening to lectures from EFL learners at different levels of proficiency, along with learning experiences and language aptitudes in order to analyze language use in more detail.

Acknowledgments

References

Brown, D. J. (1996). *Testing in language programs*. New Jersey: Prentice Hall Regents.

Canale, M., & Swain, M. (1983). From communicative competence to communicative language pedagogy. In J. Richards, & R. Schmidt (Eds.), *Language and communication* (pp. 2-27). London: Longman.

Chauncey Group International. (1998). *TOEIC technical manual*. New Jersey: Chauncey Group International.

Cronbach, L. J. (1970). *Essentials of psychological testing*. New York: Harper & Row.

Ehrman, M. E. (1996). *Understanding second language learning difficulties*. London: Sage Publications.

Eignor, D., Taylor, C., Kirsch, I., & Jamieson, J. (1998). *Development of a scale for assessing the level of computer familiarity of TOEFL examinees* (Research Reports RR98-7). New Jersey: Educational Testing Service.

Gilquin, G. (2010). Language production: A window to the mind? In H. Götzsche (Ed.), *Memory, mind and language* (pp. 89-102). Newcastle: Cambridge Scholars Publishing.

Granger, S., Dagneaux, E., Meunier, F., & Paquot, M. (2009). *International Corpus of Learner English, version*

2. Louvain-la-Neuve, Belgium: Presses Universitaires de Louvain.

Hughes, A. (2003). *Testing for language teachers.* Cambridge, UK: Cambridge University Press.

Izumi, E., Uchimoto, K., & Isahara, H. (Eds.). (2004). *Nihonjin 1200 nin no eigo spiking koopasu* [*A speaking corpus of 1200 Japanese learners of English*]. Tokyo, Japan: ALC Press (in Japanese).

Kotani, K., Yoshimi, T., Nanjo, H., & Isahara, H. (2012). Applicability of readability formulae to the measurement of sentence-level readability. In R. G. Chova, A. L. Martínez, & C. Torres (Eds.), *Proceedings of 5th International Conference of Education, Research and Innovation* (pp. 6023-6031).

Kotani, K., Ueda, S., Yoshimi, T., & Nanjo, H. (2014) A listenability measuring method for an adaptive computer-assisted language learning and teaching system. In W. Aroonmanakun, P. Boonkwan, & T. Supnithi (Eds.), *Proceedings of the 28th Pacific Asia Conference on Language, Information and Computation* (pp. 387-394).

Kirkpatrick, A. (2013). English as an international language in Asia: Implications for language education. In A. Kirkpatrick, & R. Sussex (Eds.), *English as an International Language in Asia: Implications for Language Education, Multilingual Education 1* (pp. 29-44).

Meurers, D., Ott, N., & Ziai, R. (2010). Compiling a task-based corpus for the analysis of learner language in context. In S. Featherston, & B. Stolterfoht (Eds.), *Proceedings of Linguistic Evidence 2010* (pp. 214-217).

Minematsu, N., Tomiyama, Y., Yoshimoto, K., Shimizu, K., Nakagawa, S., Dantsuji, M., & Makino S. (2003). Development of English Speech Database read by Japanese and Americans for CALT system development. *Japan Journal of Educational Technology, 27*(3), 259-272 (in Japanese, with English title and summary).

Nation, P., &Malarcher, C. (2007). *Reading for speed and fluency*. Seoul, Korea: Compass Publishing.

Nicholls, D. (2003). The Cambridge Learner Corpus: Error coding and analysis for lexicography and ELT. In D. Archer, P. Rayson, A. Wilson, & T. McEnery (Eds.), *Proceedings of the Corpus Linguistics 2003 Conference* (pp. 572-581).

Ott, N., Ziai, R., & Meurers, D. (2012). Creation and analysis of a reading comprehension exercise corpus: Towards evaluating meaning in context. In T. Schmidt, & K. Wörner (Eds.), *Multilingual corpora and multilingual corpus analysis* (pp. 47-69). Amsterdam: Benjamins. http://dx.doi.org/10.1075/hsm.14.05ott

Robb, M. P., & Gillon, G. T. (2007). Speech rates of New Zealand English- and American English-speaking children. *Advances in Speech-Language Pathology, 9*(2), 1-8. http://dx.doi.org/10.1080/14417040601013695

Yamada, M. (2015). *The role of English teaching in Modern Japan: Diversity and multiculturalism through English language education in a globalization era*. New York: Routledge.

Yasuda, K., Kitamura, K., Yamamoto, S., & Yanagida, M. (2009). Development and applications of an English learner corpus with multiple information tags. *Journal of Natural Language Processing, 16*(4), 47-64 (in Japanese, with English title and summary). http://dx.doi.org/10.5715/jnlp.16.4_47

Appendix

Table A. Example of a news article from the VOA special section

1	College prices in the United States have been rising faster than other prices for thirty years or more.
2	Recently many of the nation's top colleges have agreed to increase their financial aid.
3	But one group often has to pay the full price for college
4	This may help explain why colleges are making greater efforts to recruit them.
5	Large universities are likely to use their own representatives.
6	But smaller schools may work with independent recruiters.
7	An example is Albright College in Reading, Pennsylvania.
8	It has about one hundred foreign students, mostly from Asia.
9	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.

Table B. Example of a news article from the VOA editorial section

1	The United States and the United Nations have listed Al-Qaida in the Arabian Peninsula fugitive Fahd al-Quso as a Specially Designated Terrorist.
2	These actions will help stem the flow of finances to and inhibit the travel of this dangerous operative.
3	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.
4	The joint designation by the United States and the United Nations alerts the public that Fahd al-Quso is actively engaged in terrorism.
5	These actions," said Ambassador Benjamin, "expose and isolate individuals like al-Quso and result in denial of access to the global financial system."
6	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.
7	He was jailed in Yemen in 2002 for his part in the attack.
8	Following al-Quso's release from prison in 2007, he joined al-Qaida in Yemen.
9	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.

Table C. Example of comprehension questions for the news article

1. Which of the following is mentioned?

(a) Children are being killed, exploited and abused in ever increasing numbers in the conflict-torn country.

(b) The use of children as soldiers, torture, abduction, and other human rights abuses continue in Sri Lanka.

(c) Education is the most powerful weapon which you can use to change the world.

(d) Two decades of misrule by Saddam Hussein have taken their toll on Iraq's children.

2. According to the passage, education…

(a) is less powerful weapon to change the world.

(b) helps turn children away from violence.

(c) is merely a road to a better life.

(d) has a great chance of surviving into adulthood.

3. According to the passage, why did the number of children attending school increase?

(a) Because Nelson Mandela claimed the importance of education.

(b) Because most of the world's nations signed the Education For All initiative.

(c) Because the global economic recession threatened the gains of the past decade.

(d) Because the United States rejected universal basic education.

4. Which of the following is NOT mentioned?

(a) Education is a question of survival.

(b) Children of mothers with a full basic education are more likely to attend primary school.

(c) By 2007, the number of children attending school dropped.

(d) The United States is assisting developing countries.

5. Which of the following is NOT mentioned?

(a) Nelson Mandela was Nobel Peace Prize Laureate and human rights activist.

(b) According to the latest report, 56 million children will still be out of school in 2015.

(c) The U.S. contributed one billion dollars to Education for All.

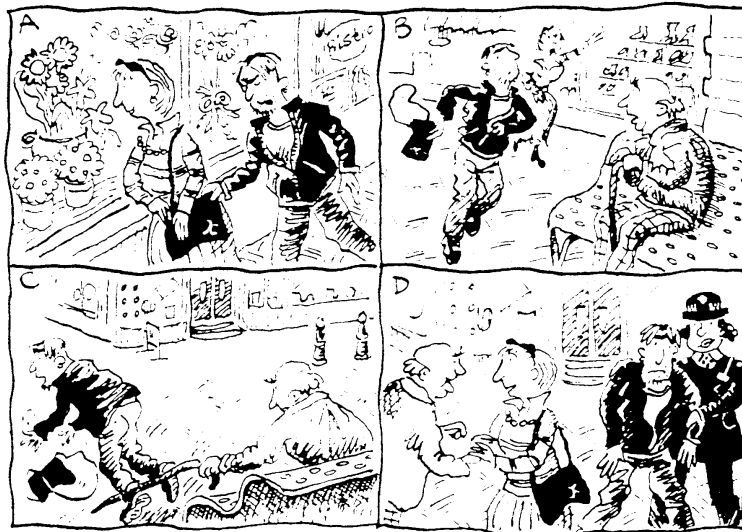(d) The United States did not assist non-governmental and multilateral organizations.

Figure A. Pictures for descriptions in the writing task

Table D. Questions for answering in the writing task

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 languages?
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
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?

Copyrights