

The Effect of English Learning in Elementary School on Students' English Language Development in Junior High School

Shigeo Uematsu

Kyoto Sangyo University, Japan

This study was an investigation of the effect of English as a Foreign Language at Elementary Schools (EFLES) in Japan on students' later English language learning in junior high school. The participants were a total of 2,000 public junior high school students in grades 7, 8, and 9 (about 220-260 students in each grade for a three-year research period). The participants received from 10-90 hours of EFLES in grades 5 and 6 of local public elementary schools. The amount of EFLES implementation basically increased every year. The research question was: To what degree does EFLES affect the students' future English skill development in vocabulary and grammar, reading, and listening measured by standardized tests administered respectively in grades 7, 8, and 9? The positive effect of EFLES was observed in the grade 7 students who had received 90 hours of EFLES classes since grade 4. An advantage was shown in higher English test scores for vocabulary and grammar, reading, and listening, though it did not reach a statistically significant level. In speaking tests conducted in grade 8, students' scores improved each year to a statistically significant level with large effect sizes ($d = 1.23$, $d = .80$ for story-telling test).

Key words: EFL, elementary school, speaking, Japanese secondary learners

INTRODUCTION

Since April 2011, foreign language activity, specifically EFLES, has been implemented nationwide by the Ministry of Education, Culture, Sports, Science and Technology (henceforth, MEXT) once a week for 5th and 6th graders. The aim of the activity is to “form the foundation of pupils’ communication abilities” and to “foster a positive attitude toward communication” (MEXT, 2009). The purpose of this paper is to investigate how EFLES will affect the English learning at a junior high school located in a *tokku* (special educational district). In this *tokku*, EFLES was implemented in grades 5 and 6 as early as 2005 and students were exposed to more English activities in each term over the following years.

The actualization of mandatory elementary school English in Japan was rather sluggish, taking a couple of decades since the Ministry of Education first set up two pilot schools in 1992. To avoid confrontation with those who argued against the implementation, English was introduced as one of the activities of the period of *sougo gakushu* (integrated studies) implementation in 2002. However, the implementation of EFLES activities increased year by year until it was adopted at 94 % of public elementary schools (Benesse, 2006). In 2008, MEXT issued an outline for uniform, once-a-week foreign language activity classes for fifth and sixth grade pupils.

Although, this kind of research approach could be conducted at any junior high school in the near future, this pilot study sheds light on how EFLES in Japan can affect students’ later English language development. Since the development of the “foundation of communicative competence” is supposedly the goal of the current elementary school English activities and has been a debatable point for many, the fiercely argued question of whether instruction at this age level is meaningful (i.e., does it lead to higher English proficiency later in secondary school?), may be partially answered by this study.

REVIEW OF THE LITERATURE

Building upon numerous EFLES studies carried out in various countries, researchers working in Japan have also conducted studies examining the effect of EFLES on later English proficiency development. They sought to investigate the long-term effect of EFLES on students' English skill development and on the transformation of their attitudes and motivation toward English language learning.

Oller and Nagato (1974) investigated the English proficiency differences between students who experienced six years of English language instruction from grade 1 in one particular school and students recruited through entrance examinations who began studying English from grade 7. The school system provided a six-year sequence of EFLES instruction in addition to a six-year EFL (English as a Foreign Language) program at the junior and senior high school levels. From grades 1 to 4, the pupils received approximately 120 minutes of English instruction per week in three 40-minute classes. In grades 5 and 6, they had two 40-minute classes per week. Only about 40 percent of the students with six years of EFLES study went on to the affiliated junior high and high school. During the first year in junior high, the EFLES students were divided into two classes based on their English proficiency. The students with and without EFLES experience used the same English textbooks, but because the EFLES students completed these textbooks quickly, two additional reading books and their accompanying workbooks were used. From the eighth grade, the EFLES and non-EFLES students were mixed indiscriminately, even though they differed in their overall proficiency levels. Three levels of cloze tests reported to be highly correlated with the TOEFL (Darnell, 1968) were administered (grade 7; $n = 104$; grade 9, $n = 89$; grade 11, $n = 103$; $N = 288$) to obtain global estimates of English proficiency for the EFLES and non-EFLES participants. The EFLES group outperformed the non-EFLES group in grades 7 and 9 to a statistically significant degree, but no statistically significant differences were found between the two groups in grade 11, suggesting that by the end of grade 11, the non-EFLES students had narrowed the gap with the EFLES students. The authors concluded that even though EFLES programs did impart some proficiency gains in the target language, there was no evidence that students

with an EFLES background would progress more rapidly than non-EFLES students in English language study at the secondary and college levels. Under certain conditions, they asserted, EFLES students would not even maintain their initial advantage over non-EFLES students in secondary school and college.

The JASTEC Kansai Project Team (2007) compared the English proficiency of 234 grade 6, 7, and 8 Japanese students in grades 6-8 from three cities in Osaka prefecture (cities A, B and C). The city A students experienced EFLES from the first grade of elementary school and had a total of 350 EFLES contact hours. The city B students started EFLES in grade 5 and had studied for a total of 70 hours. The city C students had virtually no EFLES experience; on some occasions, they heard English words such as *onion* and *fish* during activities for a class period of "international understanding". The EFLES-experienced city A and city B groups significantly outperformed the Non-EFLES city C group on the listening and interview tests. There were no differences among the three groups on the reading test. The grade 8 EFLES groups outperformed the Non-EFLES participants only on the speaking test. The researchers reported the following general findings: (a) listening skill is positively correlated with attitude toward English in grade 6, (b) all of these effects disappeared in grade 7, and (c) the correlation again emerged in both the listening and speaking domains for the advanced level in grade 8 students. However, the researchers failed to explain clearly why this last finding occurred.

Shizuka (2007) explored the relationship between English learning experiences in EFLES and English proficiency and attitude toward learning English in high school with 630 students studying in ~~one~~ a standard studies course or an English studies course at a Super English Language High school (SELHi). The students completed the ACE battery test, which was made up of listening, reading, grammar, and vocabulary components. *t*-tests revealed statistically significant differences between the Experienced and Non-experienced students on the listening ($p = .020$) and reading ($p = .001$) tests, but not on the vocabulary test ($p = .608$), grammar test ($p = .293$), or the total score ($p = .058$). Shizuka concluded that (a) Experienced students tended to be more motivated toward learning English in high school; (b) studying English for three or more years before junior high school was more effective for developing the students' overall proficiency than programs lasting for

less than three years; (c) the Experienced students' advantage over the Non-experienced students resided in their higher abilities in listening and reading skills; no advantage existed regarding vocabulary or grammar knowledge; (d) experiencing English for three years or longer was more effective for developing listening skills than doing so for less than three years; and (e) when elementary school students studied English, it did not seem to matter whether the learning took place at an elementary school or at another type of school, such as a *juku* (cram school).

As shown in the above studies, much of the research that has been conducted in the Japanese EFL context has shown that Experienced students usually outperform Non-experienced students, especially in the early stages of junior high school, in listening (JASTEC, 2007; Megumi, Yokokawa, & Miura, 1996), speaking (JASTEC, 2007; Matsukawa, 1997), reading (JASTEC, 1988), and overall proficiency (Oller & Nagato, 1974). The long-term effects of EFLES appear to exist as differences that are still observable with high school students (Megumi, Yokokawa, & Miura, 1996; Shizuka, 2007).

On the other hand, some studies have shown contradictory results. Shirahata (2002) examined the English proficiency of 115 Experienced and 122 Non-experienced junior high school freshmen in terms of phoneme perception ability, English pronunciation ability, and the amount of English output in a 5-minute self-introduction. The Experienced students were from an elementary school that was a *kenkyu* (research and development) school from 1997 to 1999. They had taken part in English activities based on *Eigo gakushu no tebiki* (The Manual for English learning) published by MEXT (2001) once a week. No statistically significant differences were found between the Experienced students and Non-experienced students on any of the test scores.

In an investigation of whether or not an EFLES approach designed to familiarize elementary school learners with English sounds through fun activities had beneficial effects on the later development of listening proficiency, Takada (2003) conducted research at a private girls' junior high school. The participants, 41 Experienced students and 49 Non-experienced students enrolled in the first year in junior high school (grade 7), completed two slightly different listening tests twice.

The participants had been studying English in junior high school for 45 and 125 hours respectively when the first and second rounds of tests were administered. She found no statistically significant differences between the Experienced and Non-experienced groups on the first-test. However, on the second test, the Non-experienced group outperformed the Experienced group ($p < .01$). She argued that the amount of exposure to English (80 hours) in the attached elementary school might not have been sufficient to put the Experienced students in an advantageous position in junior high school. She also pointed out that the Experienced students were “automatically promoted” to the junior high school while the Non-Experienced students were admitted only after surviving fierce competition to pass the entrance examination. These different screening processes might have resulted in the selection of two groups of students who differed in terms of factors such as academic achievement, motivation, and intelligence (Yukina, 2002). Taking all these conditions into consideration, one cannot simply attribute the result to the Experienced students’ “poor study habits” as Takada claimed.

It is unclear what long-term effects would be observable on learners’ subsequent linguistic development when EFLES has been implemented once a week in grades 5 and 6 from 2011, just as it was implemented in many *tokku*.

The first government-led large-scale research was conducted by the National Institute for English Policy Research (NIEPR, 2009), which investigated the immediate effects of EFLES on elementary school students in grades 5 and 6 in terms of English listening skills ($n = 3300$ per year), English speaking skills ($n = 200$ per year), their motivation to learn English, and four other issues (knowledge of the English alphabet, curriculum, instructional approach, and the relation between Japanese language proficiency and English language proficiency). The last three issues were investigated mostly through surveys and interviews obtained from local boards of education, and public elementary school administrators and teachers.

The 53 nation-wide *kenkyu* public elementary schools ($N = 3300$ for each grade) took part in the study. Tests of listening and speaking were conducted in the fall of 2007 and 2008, using nearly identical formats. They compared the test results of four groups, based on the amount of EFLES experience; Group One (1-11 hours a year), Group Two (12-22 hours a year), Group Three (23-35 hours a year), and

Group Four (36-90 hours a year) in the 2007 and 2008 academic years.

On the 21-item listening quiz (10 minutes), the students listened to audio recorded English words and short sentences classified into four categories; words, classroom English, conversation, and listening comprehension. Then, the students were asked to choose the appropriate picture from three to four options. The 2007 results and the 2008 results identified the same tendencies. Listening scores generally increased as the amount of English instruction increased, except for Group One (1-11 hours). However, only Group Four (35-90 hours) showed large score gains on the listening comprehension test.

The speaking test (approximately 5 minutes) conducted at three schools consisted of 31 questions in 2007 and 28 items in 2008. Questions were based on the thirteen pictures and the same format was used for grades 5 and 6. In the 2008 version, some items were replaced to eliminate items that were not responded to. They reported that both in 2007 and 2008, the grade 6 students' responses were more positive than those of the grade 5 students; hence the grade 6 students' speaking scores had improved.

In the wider context of EFL, the studies of early EFLES designed to investigate the effect of the age of onset broadly present a consistent picture of subsequent skill development in the target language. Focusing on the age of onset, Cenoz (2003) administered comprehensive English tests measuring oral production, reading comprehension, listening comprehension, and writing skills in order to investigate the effect of age on general English proficiency. The results indicated that the older learners obtained significantly higher results than the younger learners on most of the measures of English proficiency. Likewise, the studies of Cenoz (2009), Lecumberri and Gallard (2003), and Mora (2006) seem to be generally in line with the statement of Krashen, Long, and Scarcella (1979), who distinguished between the rate of acquisition (short-term attainment) and ultimate attainment (long-term attainment) from various immigrant studies in the field of SLA, and then asserted that older learners have an advantage where short-term achievement is concerned (Long, 1990; Scovel, 1988; Singleton & Ryan, 2004). In regards to this issue, Munoz (2006) summed up the possible educational implications for EFL context:

In sum, these findings suggest that second language success in a foreign

language context may be as much as a function of exposure as of age. Exposure needs to be intense and to provide an adequate model. Initial age of learning seems more relevant for skills that can be acquired implicitly, whereas age at learning can be seen as a factor explaining the rate of learning of most skills (p. 34).

To welcome the MEXT's decision as the first step toward the fulfillment of English education in a new direction, the purpose of this study was to investigate whether or not there was any effectiveness of EFLES on learners' later development in terms of English language proficiency when the age of onset was lowered and the amount of exposure increased in the case of once-a-week instruction in Japanese context. Is it possible to find any difference when the amount of exposure was limited to small numbers and the age of onset ranged from grade 4 to grade 6? In order to answer these inquiries, the research question is as follows: To what degree do the 2007, 2008, 2009 participants differ in terms of their English listening, speaking, reading, vocabulary, and grammar skills?

THE STUDY

Participants

The participants in this study were a total of 2000 students from Junior High School "A" in western Japan. The school is a co-educational public school offering English instruction four times a week to an approximately equal number of male and female students. A native speaker of English Assistant Language Teacher (ALT) occasionally helped the Japanese English teacher conduct the lessons. One of the four School A English classes was called *kokusai komyunikeishon* (International Communication, henceforth IC) class, the continuation of EFLES activities, and was focused on the development of the students' speaking and listening skills. This course was started from grade 5 in 2005 and was lowered to grade 1 in 2006.

About 220-260 students in grades 1 through 3 in A junior high school took part in

the study each year from the 2007 through the 2009 academic year¹. The amount of EFLES experience differed at the time of research administration. For example, in the 2007 academic year, the first-year (grade 7) students experienced 35 hours of EFLES instruction in grades 5 and 6 (70 total contact hours) prior to entering junior high school, while the second-year (grade 8) students experienced 35 hours of EFLES instruction in grade 6, and the third-year (grade 9) students received less (approximately ten hours).

The participants in School A came from two elementary schools; the curricula at the two schools are supposed to be similar in order to provide a uniform quality of English instruction and avoid any perception of unfairness. English is taught in these elementary schools based on the unified IC curriculum adopted by the city, which is based on a graded series of commercial English supplementary textbooks, *Minna no Eigo* (Ito & Hasegawa, 2002).

After removing five participants who (a) failed to sit for one or more of the proficiency test components, (b) had lived in an English-speaking country more than one year, or (c) answered with a single response (e.g., all 1s), approximately, 200-220 participants were retained.

Instrumentation

English Proficiency Test

The JACE (Junior High School Assessment of Communicative English) battery Level 1 through Level 3 English proficiency tests were administered to the grade 7, grade 8 and grade 9 students in March of 2008, 2009, and 2010. The JACE test (Levels 1-3), which is published by ELPA (Association for English Language Proficiency Assessment), was designed to measure English proficiency of Japanese junior high school students. The test is made up of three subsections: (a) vocabulary and grammar (22 items in 2 sections; maximum 100 points), (b) reading (10 items in 3 sections; maximum 100 points), and (c) listening (18 items in 2 sections; maximum 100 points). According to ELPA, the test items are standardized using item response theory. The reliability estimates for the Level 1, Level 2 and Level 3 test scores are .81, .84 and .82, respectively.

Oral Interview Test

The author and the school agreed that participants from one class in grade 8 took part in a ten-minute interview test in March of 2008, 2009, and 2010. The interview test was made up of three parts: (a) four questions and answers related to the student's self-introduction, (b) four questions and answers based on a picture, and (c) a story-telling task based on the picture. The same task format with the same sets of questions and picture were used each year. The conversation section, parts (a) and (b), was made by referencing the STEP *Eiken* 3rd grade conversation test. The pictures used in the story-telling task were drawn by college students. Two Japanese university English teachers, including myself, conducted the interviews in different rooms at School A. Prior to the interview, the two raters met and talked about the interview tests and the assessment criteria. The two raters assessed the candidates independently as they worked in two separate rooms during a day of interviews. All interviews were videotaped and one outside rater, a native speaker of English, watched the videotapes and rated the students' performances. Each Japanese rater watched the videotapes of the students assessed by the other Japanese interviewer and provided ratings for those students. The raters evaluated the performances of the students on the items designed to measure their oral English proficiency. The evaluation sheet was based on the STEP *Eiken* test rating sheet. The two parts of the tasks (conversation and story-telling) were rated on a 5-point Likert scale (1 = Poor; 2 = Not good; 3 = Acceptable; 4 = Good; 5 = Excellent) in terms of three criteria: (a) Attitudes toward communication and willingness to communicate; (b) Fluency, intonation, rhythm, and pronunciation; and (c) Vocabulary, grammar, and word usage.

DATA ANALYSIS

The data analysis procedures were as follows. There are two dependent variables (DVs) for each grade: The first linguistic variable consisted of three JACE battery subtest scores for the vocabulary and grammar, listening comprehension, and reading comprehension subtests. The second linguistic variable is the speaking

measure, which is made up of the conversation and story-telling sections. The independent variable is academic year (2007, 2008, 2009).

The research question, “To what degree do the 2007, 2008, 2009 participants differ in terms of their English listening, speaking, reading, vocabulary, and grammar skills?” was addressed by conducting a series of the one-way ANOVA tests of the JACE test scores and the Rasch person ability measures from the interview test scores as dependent variables. Following Tabachnick and Fidell (2006), the assumptions for the one-way ANOVA were checked in terms of (a) normality of sampling distribution of means, (b) independence of errors, (c) homogeneity of variance, and (d) absence of outliers.

The interview test data were checked using FACETS 3.62 for Windows (Linacre, 2008). The FACETS program, which implements the multi-faceted Rasch model (Linacre, 2008), incorporates an algorithm that “expresses the probabilistic expectations of item and person performances when one construct is held to underlie the developmental sequence by the observation schedule” (Wright & Stone, 1979). All raters showed acceptable fit to the Rasch model. Rater 1 was the most severe rater with an estimate of .43, while Rater 3, the most lenient rater, had an estimate of .02. All items displayed a good fit to the Rasch model. The inter-rater reliability was .84 for the conversation task and .83 for the story-telling task. Since there were no data flaws, the raw scores were used for the ANOVA analysis.

RESULTS

JACE Test Scores

Descriptive statistics for the grade 7 JACE tests are shown in Table 1. The results follow the table.

TABLE 1
Descriptive Statistics for Grade 7 JACE Scores

	HE	<i>N</i>	<i>M</i>	<i>SD</i>	SE	95% CI	Min	Max	
VG	2007	70	209	56.59	13.75	.95	[54.71, 58.46]	25	100
	2008	70	198	54.68	15.89	1.12	[52.45, 56.90]	0	100
	2009	90	199	56.82	11.16	.79	[55.26, 58.38]	25	100
R	2007	70	209	55.97	19.07	1.31	[53.37, 58.57]	0	100
	2008	70	198	58.83	21.14	1.50	[55.87, 61.80]	0	100
	2009	90	199	60.99	14.34	1.01	[58.99, 63.00]	10	90
L	2007	70	209	58.75	10.88	.75	[57.27, 60.24]	30	90
	2008	70	198	58.80	12.45	.885	[57.06, 60.55]	0	100
	2009	90	199	60.20	20.26	1.43	[57.37, 63.04]	0	100

Note. VG (vocabulary and grammar scores), R (reading scores), L (listening scores), HE (hours of EFLES) ; CI = confidence interval

For vocabulary and grammar scores, the result of one-way ANOVA, with *Scheffe* multiple comparisons revealed there was no statistically significant difference $F(2, 603) = 1.461, p = .233$). However, 2009 students (90 hours' EFLES from grade 4) slightly outperformed 2008 students ($d = .12$) who had less EFLES experience (70 hours from grade 5).

For reading scores, the result of ANOVA showed there was a statistically significant difference $F(2, 603) = 3.833, p = .022$). 2009 students slightly outperformed 2008 students ($d = .12$), and outperformed 2007 students ($d = .28$) to a statistically significant level.

For listening scores, the result of one-way ANOVA revealed there was no statistically significant difference $F(2, 603) = .602, p = .548$). However, 2009 students slightly outperformed 2008 students ($d = .10$). Descriptive statistics for the grade 8 JACE tests are shown in Table 2, followed by the results.

TABLE 2
Descriptive Statistics for Grade 8 JACE Scores

	HE	<i>N</i>	<i>M</i>	<i>SD</i>	SE	95% CI	Min	Max	
VG	2007	35	171	51.67	14.66	1.12	[49.46, 53.89]	20	171
	2008	70	210	52.48	14.20	.98	[50.55, 54.41]	20	210
	2009	70	204	51.27	15.16	1.06	[49.18, 53.36]	0	204
R	2007	35	171	47.21	23.57	1.80	[43.65, 50.77]	0	171
	2008	70	210	48.36	22.50	1.55	[45.30, 51.42]	0	210
	2009	70	204	48.80	24.40	1.70	[45.43, 52.17]	0	204
L	2007	35	171	51.25	14.02	1.07	[49.13, 53.36]	20	171
	2008	70	210	52.29	11.84	.81	[50.67, 53.90]	0	210
	2009	70	204	51.03	12.68	.88	[49.28, 52.79]	0	204

Note. VG (vocabulary and grammar scores), R (reading scores), L (listening scores), HE (hours of EFLES) ; CI = confidence interval

For vocabulary and grammar scores, the result of ANOVA revealed there was no statistically significant difference $F(2, 582) = .365$, $df = 2/582$, $p = .694$). 2008 students slightly outperformed 2007 and 2009 ($d = .07$) students.

For reading scores, the result of ANOVA showed there was no statistically significant difference $F(2, 582) = .224$, $p = .80$). 2009 and 2008 students slightly outperformed 2007 students ($d = .07$, $.06$ respectively).

For listening scores, the result of ANOVA revealed there was no statistically significant difference $F(2, 582) = .560$, $p = .571$). 2008 students slightly outperformed 2009 and 2007 students ($d = .05$, $.11$ respectively). Descriptive statistics for the grade 9 JACE tests are shown in Table 3, followed by the results.

TABLE 3
Descriptive statistics for Grade 9 JACE scores

	HE	<i>N</i>	<i>M</i>	<i>SD</i>	SE	95% CI	Min	Max	
VG	2007	10	200	67.71	20.97	1.48	[64.78, 70.63]	20	100
	2008	35	180	65.00	22.02	1.64	[61.76, 68.24]	25	100
	2009	70	222	65.00	23.00	1.54	[61.96, 68.04]	10	100
R	2007	10	200	71.63	25.36	1.79	[68.09, 75.17]	0	100
	2008	35	180	69.33	21.08	1.57	[66.23, 72.44]	10	100
	2009	70	222	67.97	24.68	1.65	[64.71, 71.24]	10	100
	2007	10	200	71.18	14.58	1.03	[69.14, 73.21]	30	100

L	2008	35	180	70.48	16.02	1.19	[68.13, 72.84]	35	100
	2009	70	222	71.53	15.78	1.05	[69.44, 73.61]	25	100

Note. VG (vocabulary and grammar scores), R (reading scores), L (listening scores), HE (hours of EFLES); CI = confidence interval

For vocabulary and grammar scores, the result of ANOVA revealed there was no statistically significant difference $F(2, 599) = 1.004, p = .367$. 2007 students slightly outperformed 2008 and 2009 ($d = .10$) students.

For reading scores, the result of ANOVA showed there was no statistically significant difference $F(2, 599) = 1.247, p = .288$. 2007 students slightly outperformed 2008 and 2009 students ($d = .11, .17$ respectively).

For listening scores, the result of ANOVA revealed there was no statistically significant difference $F(2, 599) = .230, p = .795$. 2009 students slightly outperformed 2008 and 2007 students ($d = .07, .03$ respectively).

Interview Test Scores

For the respective interview test scores, the result of one-way ANOVA, with *Scheffe* multiple comparisons were employed for further analysis. Since we administered the interview test on only one class in grade 8, the N-size was small (31 for 2007, 35 for 2008, 35 for 2009). Descriptive statistics for the grade 8 JACE tests are shown in Table 4, followed by the results.

TABLE 4
Descriptive Statistics for Interview Test Scores: Conversation Task

	HE	N	M	SD	SE	95% CI	Min	Max	
Conversation task	2007	35	31	11.54	2.62	.47	[10.58, 12.50]	3	15
	2008	70	35	13.55	1.65	.28	[12.99, 14.12]	3	15
	2009	70	35	14.03	1.18	.20	[13.62, 14.43]	3	15

Note. HE (hours of EFLES); CI = confidence interval

For the conversation task scores, the result of one-way ANOVA revealed there was a statistically significant difference $F(2, 98) = 16.01, p = .000$ for the conversation task scores. According to the *Scheffe* test, 2007 students ($M = 11.54$,

$SD = 2.62$) were outperformed by 2008 students ($M = 13.55$, $SD = 1.65$); $t(64) = -3.79$, $p = .000$, $d = .75$ (close to large effect size) 2008 students ($M = 13.55$, $SD = 1.65$) students were also outperformed by 2009 students ($M = 14.03$, $SD = 1.18$); $t(68) = -.89$, $p = .03$, $d = .22$. As for the story-telling task scores, the following results were obtained. Descriptive statistics for the grade 8 JACE tests are shown in Table 5, followed by the results.

TABLE 5
Descriptive statistics for Interview test scores: Story-telling task

	HE	N	M	SD	SE	95% CI	Min	Max
Story-telling task	2007	35	7.57	3.10	.56	[6.43, 8.71]	3	15
	2008	70	11.40	2.93	.49	[10.40, 12.41]	3	15
	2009	70	13.34	1.39	.24	[12.86, 18.82]	3	15

Note. HE (hours of EFLES); CI = confidence interval

The result of one-way ANOVA revealed there was a statistically significant difference $F(2, 98) = 42.67$, $p = .000$ for the story-telling task scores. According to the *Scheffe* test, 2007 students ($M = 7.57$, $SD = 3.10$) were outperformed by 2008 students ($M = 11.39$, $SD = 2.93$; $t(64) = -5.06$, $p = .000$, $d = 1.23$ (large effect size). 2008 students ($M = 11.39$, $SD = 2.93$) were also outperformed by 2009 ($M = 13.34$, $SD = 1.39$) $t(68) = -3.91$, $p = .000$, $d = .80$ (large effect size).

DISCUSSION

In this section, I discuss the results of the study, interpret the findings for each analysis, and compare and contrast them with the findings of previous studies.

Grade 7, 2009 students who had received 90 hours' EFLES from grade 4 slightly outperformed 2008 students ($d = .12$) for vocabulary and grammar scores. For reading scores, they also slightly outperformed 2008 students ($d = .12$), and outperformed 2007 students ($d = .28$) students to a statistically significant level. For listening scores, they slightly outperformed 2008 students ($d = .10$). Both 2007 and 2008 students had received 70 hours' EFLES from grade 5. As a result, 2009

students who started EFLES from grade 4 performed best in all three JACE linguistic test scores. These results were in line with the studies in which EFLES was reported to exert positive effects on overall proficiency (Oller & Nagato, 1974), and listening proficiency in the early stages of junior high school (JASTEC, 2007; Megumi, Yokokawa, & Miura, 1996). The results also support the previous findings concerning the subsequent positive effect of FLES on learners' listening skills outside of the Japanese context (Burstall, 1975; Kwon, 2006), hence also support "the earlier, the better" notion in Second Language Learning, which has been explained by the Critical Period Hypothesis, stating that the age of onset for implicit language learning is the significant factor (Aoyama et al., 2008; Lenneberg, 1967; Long, 1990; Scovel, 2006).

However, it should be noted that the 2009 students barely outperformed the previous cohorts, even though the N sizes for each grade were reasonably large (around 200 per year). The most convincing reason might be that the amount of exposure was not sufficient, only 90 hours in total, 20 hours more than the previous cohorts who started in grade 5. If 2009 students had had a larger amount of exposure, the difference could have been greater.

As far as grade 8 students were concerned, 2009 students failed to show their advantage over the previous 2008 and 2007 students in JACE test scores. The amount of exposure to English (70 hours) in elementary schools might not have been sufficient to put them in an advantageous position in grade 8, when compared with the previous cohorts who had received the same exposure (70 hours) or less (35 hours).

However, in speaking tests, both for conversation test task scores and story-telling task scores, 2009 students outperformed the 2008 and 2007 students to a statistically significant level. Furthermore, each year students' scores improved to a statistically significant level both for conversation task scores and story-telling task scores with large effect sizes. The possible reasons for these are a) students continued the IC curriculum, the focus of which was on developing students' communicative skills, in junior high school once a week, b) more students were prepared for the interview conducted in English because the city subsidized the Step Eiken test fee, aiming to increase the number of students who could obtain the Step

certificates, and c) the teacher might have been better prepared for the interview year by year, hence might have guided the students in how they could successfully sit for the interview. In any case, the results go with previous findings concerning the positive effect of EFLES on speaking skills, especially in the early stages of junior high school (JASTEC, 2007; Matsukawa, 1997; Shinohara, 1999). Matsukawa (1997) stated, based on her story-telling task results, that EFLES could have played an important role in fostering effective communication strategies, such as responding as quickly as possible, and that the EFLES experienced students generally produced a greater amount of output.

Cenoz (2003) contended the advantage of an early exposure group was “related to educational factors and particularly to input and teaching methods used in secondary school classes when an oral-based approach and a very active methodology based on drama and story-telling was used” (p. 90). Here, EFLES could exert a strong effect on forming the foundation of “communication,” if a communication-based class is continued in junior high school.

For grade 9 students, 2009 students who had received 70 hours of EFLES from grade 5 failed to outperform both 2008 students (35 hours from grade 6) and 2007 students (only 10 hours, occasionally). Instead, 2009 students were slightly outperformed by 2007 students in all of the JACE test scores. These results overlap with the findings of Oller & Nagato (1974) that stated EFLES students would not even maintain their initial advantage. Even though they had received more EFLES, the other groups easily caught up. In grade 9, the least amount cohort (10 hours) outperformed the 35-hour cohort and the 70-hour cohort for vocabulary and grammar scores and in reading scores.

CONCLUSION

Here, I briefly summarize the main findings of the study, discuss the limitations of the study, and make suggestions for future research. First, speaking test scores for 2007, 2008, 2009 participants improved to statistically significant levels each year as the amount of exposure increased in grade 8. EFLES can exert a powerful effect

on fostering the foundation of communication skills in English when an English class focusing on communication is continued in junior high school. Second, the 2009 cohort which started the EFLES program earliest, from grade 4, was the most successful in developing their overall English proficiency measured by JACE test scores for vocabulary and grammar, reading, and listening when compared with 2008 and 2007 students, but they did not reach a statistically significant level. Third, these advantages, however, might not last long, maybe around a year. In grade 8 and grade 9, 2009 participants were outperformed by 2007 and 2008 students. The effect of once-a-week EFLES in grade 5 and grade 6 classes might have yielded to junior high school English teaching, except for speaking.

Although this study provides important results to the field of EFLES and EFL, several limitations affect the findings of the study. The first limitation of this study concerns the sample size of the interview test. The relatively small sample negatively affected the generalizability of the results. The small sample size also limited the study by not permitting the investigation of subgroups. The second limitation occurred because the interview tests used in this study were based on the STEP (*Eiken*) third grade test, and the construct validity of the STEP test is questionable because of insufficient information indicating this interview test is a valid index of students' English speaking proficiency.

After addressing the above-mentioned limitations, a series of future studies could be carried out. The following research questions could be investigated in the future. (a) How do the nature of English instruction at the respective elementary schools and teacher variables influence the learners' future linguistic development? (b) How do the participants' motivational and attitudinal variables change over time? (c) How do the nature of English instruction and teacher variables at junior high school influence the students' EFLES experience? The use of multiple approaches for data collection can reduce the risks of making incorrect generalization based on limited information. Thus, adding qualitative approaches to data collection and thereby using a mixed-methods approach would be one way to gain a better understanding of EFLES in Japan.

NOTES

1. This study was originally designed to continue for four years from 2007 through the 2010 academic year (ending in March, 2011). However, the 2010 test administration was not able to be conducted in January, 2011 for various reasons.

THE AUTHOR

Shigeo Uematsu holds an M.A. from Columbia University and Ed.D. from Temple University, and is currently a professor in the Dept. of Cultural Studies of Kyoto Sangyo University. His recent publications include *The Effect of English Learning in Elementary Schools on Students' English Language Skills and Their Affective Variables in Junior High Schools* (2010) and *The Long-term Effectiveness of English Language Instruction at Japanese Elementary Schools* (2009).

Email: uematsu@cc.kyoto-su.ac.jp

ACKNOWLEDGEMENTS

The research reported in this paper was supported by Grant-on Aid for Scientific Research (C) (19520530, 23520769) from the Ministry of Education, Culture, Sports, Science and Technology. I would like to thank the three anonymous Asia TEFL reviewers for their comments on earlier version of this article.

REFERENCES

- Aoyama, K., Guion, S. G., Flege, J. E., Yamada T., & Akahane-Yamada, R. (2008). The first years in an L2-speaking environment: A comparison of Japanese children and adults learning American English. *International Review of Applied Linguistics in Language Teaching (IRAL)*, 46, 61-90.
- Benesse (2006). *Dai nikai shougakkou eigo ni kansuru kihon chosa* [The Second Basic Survey on EFLES]. Retrieved May 17, 2012 from <http://benesse.jp/berd/center/open/>

report/syo_eigo/2010_dai/dai_03.html

- Burstall, C. (1975). Primary French in the balance. *Educational Research*, 17, 193-197.
- Cenoz, J. (2003). The influence of age on the acquisition of English: General proficiency, attitudes and code mixing. In M. P. Garcia Mayo & M. L. Garcia Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 77-93). Clevedon, England: Multilingual Matters.
- Cenoz, J. (2009) *Towards Multilingual Education: Basque Educational Research in International Perspective*. Clevedon, England: Multilingual Matters.
- Darnell, D. K. (1968). The development of an English language proficiency test using a closentropy procedure. *ERIC ED* 024-039.
- Ito, Y., Hasegawa, J. (2002). *Minna no eigo*. Book I, II, III. Tokyo, Japan: Gakken.
- JASTEC Project Team (1988). Soki eigo gakushusha no tsuiseki chousa: Dai san pou [A follow-up study on the effects of early English learning: Report III]. *JASTEC Journal*, 7, 43-63.
- JASTEC Kansai Project Team (2007). Shougakko eigo gakushu keikensha no tsuiseki chosa to shouchugakkou eigo kyouiku eno shisa [A follow-up study of EFLES experienced students and its implication on the future English instruction at elementary school and junior high school]. *Kinki University Language Education Bulletin*, 7(2), 123-180.
- Krashen, S., Long, M., & Scarcella, R. (1979) Age, rate and eventual attainment in second language acquisition. *TESOL Quarterly*, 13, 573-578.
- Kwon, O. (2006) Impacts and effects of ten years of elementary school English education in Korea. *GTEC 2006 Report*, 78-85.
- Lecumberri, M. L. G., & Gallardo, F. (2003). English FL sounds in school learners of different ages. In M. P. Garcia Mayo, & M. L. Garcia Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 115-135). Clevedon, England: Multilingual Matters.
- Lenneberg, E. H. (1967). *Biological Foundations of Language*. New York, NY: Wiley
- Linacre, J. M. (2008). *WINSTEPS: Rasch-model computer programs, version 3.64*. Chicago, IL: Winsteps.com.
- Long, M. (1990). Maturation constraints on language development. *Studies in Second Language Acquisition*, 12, 251-285.
- Matsukawa, R. (1997). *Shogakko ni eigo ga yatte kita*. [English visits elementary schools]. Tokyo, Japan: Apurikotto.
- Megumi, T., Yokokawa, H., & Miura, I. (1996). Soki eigo gakushu keikensha no chukou ni okeru seiseki [High school grades in early learners of English]. *JASTEC Journal*, 15, 27-35.
- MEXT (2001). *Practical handbook of elementary school English activities*. Tokyo, Japan: Kairyudo.
- MEXT (2009). *Shogakko gakushu shido yoryo kaisetsu. Gaikokugo katsudo hen* [New

- government guidelines for teaching elementary school. Chapter of foreign language activities]. Retrieved May 17, 2012 from http://www.mext.go.jp/component/a_menu/education/micro_detail/_icsFiles/afieldfile/2009/06/16/1234931_012.pdf
- Mora, J. C. (2006). Age effect on oral fluency development. In C. Munoz (Ed.), *Age and the rate of foreign language learning* (pp. 65-88). Clevedon, England: Multilingual Matters.
- Munoz, C. (2006). The effects of age on foreign language learning: The BAF project. In Carmen Munoz (Ed.), *Age and the rate of foreign language learning* (pp. 1-40). Clevedon, England: Multilingual Matters.
- NIEPR. (2009). *Shogakko ni okeru eigo kyoiku no arikata ni kansuru chousa kenkyu: kenkyu hokoku kai siryo*. [A Research investigating the ideal nature of English education at elementary school: A summary report]. Tokyo, Japan: National Institute for English Policy Research
- Oller, J., & Nagato, N. (1974). The long-term effects of FLES: An experiment. *The Modern Language Journal*, 58, 15-19.
- Scovel, T. (2006). Age, Acquisition, and Accent. *Linguistic Insights - Studies in Language and Communication*, 22, 31-48.
- Shinohara, Y. (1999). Souki eigo kyouiku ga hatsuon ni oyobosu koka nit suite. [A study of the effect of EFLES on the pronunciation]. *JASTEC Journal*, 18, 1-22.
- Shirahata, T. (2002). Kenkyu kaihatsukou de eigo ni sesshita jidou no eigo nouryoku chosa. [A research investigating English proficiency of the students who were exposed to English at a kenkyu elementary school]. *Shizuoka University Educational Bulletin*, 33, 195-215
- Shizuka, T. (2007). Effects of learning English in elementary school days on the proficiency of, and the attitude toward, the language in high school years. *JACET Journal*, 45, 31-45.
- Singleton, D., & Ryan, L. (2004). *Second language acquisition: The age factor* (2nd ed.). Clevedon, England: Multilingual Matters.
- Tabachnick, B., & Fidell, L. (2006). *Experimental designs using ANOVA*. New York, NY: Duxbury.
- Takada, T. (2003). Souki eigo kyouiku keikensha to mikeikensha no chukangengo no bunseki. [An analysis of interlanguage of those who have experienced English learning at elementary schools and those who did not.] *STEP Bulletin*, 15, 159-170.
- Wright, B. D., & Stone, M. (1979). *Best test design*. Chicago, IL: MESA Press.
- Yukina, K. (2002). Eigo kishusha to eigo mishusha no chuggako deno eigoryoku to chino no soukan: shogakko dankai deno eigo kyoiku no hitsuyo sei. [The correlation between EFLES Experienced students and Non-experienced students at the junior high school stage: The investigation of the necessity of EFLES]. *Waseda Kyoiku Hyoron*, 15(1), 95-109.

